

Econophysicists matter

Economics and physics are two disciplines that, contrary to widespread perceptions, have significant common agendas. Shame, then, that the professionals don't do more to recognize the fact.

After hearing a talk on the application of physics to the social sciences, a physicist in a notoriously traditional department was heard to mutter that it was all very well but it wasn't 'real physics'. It was an article of faith to him that many-body theories in physics could not be applied to animate objects.

Now that it seems clear that bacteria, locusts and even road traffic undergo types of dynamic phase transition, this objection is hard to sustain. But the idea that physics can tell us something about a human system as complex as the economy — that there exists a kind of econophysics — seems harder to swallow.

Would-be econophysicists can therefore find themselves damned from both directions: physicists don't think of the field as physics, whereas economists don't recognize it as their discipline either. Acceptance by the economics community seems a particularly long way off: even fully fledged economists are ostracized by the mainstream if they do not embrace the tenets of 'neoclassical' economic theory, no matter how untenable its principles — identical, utility-maximizing economic agents operating in an equilibrium market — now seem.

The refusal of the economic mainstream to engage with econophysics is lamentable and makes it difficult for physicists to recognize and learn from their mistakes. But it is going to be a fact of life for some time to come. Nevertheless, there is now a body of respected economists who acknowledge the potential value of ideas and tools taken from other sciences, including physics, and who are receptive to the efforts of physicists.

As the News Feature on page 686 illustrates, this opportunity to build links with receptive economists should not be squandered by a lack of quality control in econophysics. Current standards in the field are extremely variable; there is sometimes a sense that physicists are content to find a vaguely plausible way of mapping

some economic question onto a familiar physics-based model, to characterize its behaviour and plot a phase diagram, and leave the matter at that.

This is not particularly good physics; it is certainly poor social science; and it may prove irrelevant to the questions that really matter in understanding economic behaviour. Journals that are willing to publish econophysics must be more vigilant and thorough in their review procedures, and be willing to seek out sympathetically minded economists (who do exist) for advice.

But as some econophysicists argue, it will be hard to improve standards while their efforts are necessarily a sideline tolerated only as long as they also work on 'real' physics (for example, there is not a single chair of econophysics anywhere in the world). This encourages not only the perception but also the reality of a certain dilettantism in the field.

Economics cannot be something for physicists to dabble with. Indeed, the challenge it poses is in some ways even more daunting than that facing physicists who wish to work on biological problems — at least in biology there is a core body of knowledge that, however complex, represents more or less a consensus view. In economics it seems likely that the fundamental principles have yet to be defined, whatever neoclassical theorists say. That is why no single textbook will bring physicists up to speed, and why they have some serious homework to do. They need to be given time to do it.

Econophysics has yet to assemble the critical mass needed to be self-sustaining. If it is not to become extinct, it needs to be made a more attractive career option than it currently is. ■

"The refusal of the economic mainstream to engage with econophysics is lamentable and makes it difficult for physicists to learn from their mistakes."

Brute force in the clinic

Clinical microbiologists should catch up with their colleagues and use metagenomics.

You will not die alone, but with a trillion-strong audience of bacteria and other microscopic buddies. They are glued to our tongues, swarming in our intestines and hitching a ride in our noses. They aid our digestion and shape our immune systems.

Delving into these orifices and their microbial residents is exciting many microbiologists — a sentiment apparent in certain sessions of the American Society for Microbiology meeting in Orlando late last month. Keen to document exactly which microbes are there, many researchers are using brute-force 'metagenomic' studies in

which they extract microbial DNA from the body cavity of choice and hurl it all into a sequencing machine. They can identify hundreds or thousands of microbial genes in one swoop.

The discussions in adjacent rooms offered a sharp contrast. Clinical microbiologists — those tasked with identifying the meaner, infectious microbes that invade our bodies and hospitals — use a different set of time-honoured techniques.

Conventionally, they take a swab or sample and attempt to identify the culprit microbe by growing it on certain agar media or examining the shape of a fungi's tentacles under the microscope. Genetic tests are not routine and are typically done to identify one organism at a time. Results can take a day or two to come through. Doctors facing a severe infection use pre-emptive antimicrobial drugs that might later prove fruitless.

Clinical microbiologists have good reasons for sticking with these



seemingly old-fashioned techniques. Above all, they are foolproof, cheap and definitive — why use an expensive sequencing machine when a culture plate will do? Genetic tests run a high risk of artefacts because one stray contaminating bacterium could produce a false result — and clinical microbiologists have only one precious sample and cannot tolerate errors when a patient's life is at stake.

But used alongside culture methods, the potential rewards of sophisticated genetic analyses are tantalizing. They can reveal a host of bacteria that we cannot grow in culture. Some researchers are designing gene chips that could, if adapted for use in hospitals, reveal which of a range of microbes is there, what strain it is, and which antibiotic-resistance genes and virulence factors it harbours. Used widely, such profiling could help show where a microbe has come from and whether it is spreading.

There is gap to be bridged here. Basic researchers have created ever more crafty ways to hunt down microbes. Clinical microbiologists, who have to deal with hundreds of different diagnostic tests simultaneously, want robust, high-throughput methods that have clear benefits in helping them choose a drug or curtail an infection. As is so often the case, research that converts the former into the latter is in short supply — and neither group has the time or expertise to do it.

More cash for this type of research would, as always, help. It could be used to train more clinical microbiologists in advanced molecular techniques or equip laboratories to trial new diagnostic methods. Meanwhile, basic and clinical researchers can make a simple start, by talking to each other a little more, rather than sitting in separate conference rooms.

Genetic snapshots of the teeming mass of bodily microbes complicate things further. Take, for example, a study released last month into the bane of cystic fibrosis sufferers, *Pseudomonas aeruginosa* (E. E. Smith *et al. Proc. Natl Acad. Sci. USA* **103**, 8487–8492; 2006). By tracking one infection over eight years, the study showed that this bacterium is a moving target: at any one time, many different lineages are present, presumably nestled into different crevices of the lungs. In order to understand and halt an infection, it may not be enough to know that *P. aeruginosa* is there — we might need to know all the different variants and how they interact with each other and with other microbes. This area is ripe for research and might reveal a chink in the microbe's antibiotic-resistant armour.

Metagenomic studies are fascinating for those of us grown-ups still childishly interested in our own dank, dark recesses. But microbial surveys have urgent application in medical care. The research and medical community must ensure they are used there. ■

Peer review on trial

Introducing two *Nature* initiatives.

This publication champions the value of editorially driven or editorially selected content. But we are always keen to try new things, and we are now experimenting around the edges of that principle, to make the most of online interactivity.

In recent months we have started attaching blogs to every daily news item on our news@nature.com website; see, for example, a debate on the 'missing link' *Tiktaalik roseae* (http://blogs.nature.com/news/blog/2006/04/the_fish_that_crawled_out_of_t.html).

Blogs are unlikely to replace journalism, but are probably here to stay as a valuable complement. Less certain is the outcome of a trial that we launch this week: a test of a particular type of open peer review. The trial is accompanied by a general online debate about peer review; see www.nature.com/nature/peerreview/index.html.

During the trial, which will last several months, *Nature's* traditional approach to peer review will continue: typically, we send selected submissions to two or three experts whose identities are kept confidential. We believe that this approach works well. Meanwhile, over the next few weeks, the web debate will explore other approaches, as well as the potential for online techniques to unpack the various functions of conventional journals, the ethics of peer review, and more.

Our online trial opens up a parallel track of peer review for submitted papers for authors willing to go down that route. The traditional process will still be applied to all submissions selected for peer review. But we will also offer to post the submitted manuscript onto an open website. Anyone can then respond to it by posting online comments, provided they are willing to sign them. Once

Nature's editors have received all the comments from their solicited confidential reviewers, the open website will cease to take comments, and all the opinions will be considered by the editors as well as the authors.

The open procedure will only be applied to the first version of a manuscript. Once a manuscript has been accepted or rejected, the editor will privately assess the value of each open contribution.

The trial will run for several months. Once all final decisions on the relevant manuscripts have been reached, we will conduct an aggregated assessment of the comments received as well as the work required to do justice to them. We will report the result in these pages, and consider whether our peer-review procedures should be adapted appropriately. We expect any changes to be complementary to our existing processes, not a replacement.

We use the word 'trial' rather than 'experiment' advisedly. An experiment would set out to establish the fundamental benefits and disadvantages of this approach to open peer review. We do not claim such grand ambitions. We simply want to suck it and see, and a survey of recent authors has demonstrated enough interest in the idea to make such a trial worthwhile.

Any paper that we post on our open website for comments will of course be fair game for public access and for journalists. Prior media coverage will not endanger their acceptance. But journalists will be aware that papers not subjected to peer review carry their own risks.

Here, as in other ways, *Nature* and its publishers are exploring and expanding our opportunities online. But our core goal remains as always: to bring our readers the most stimulating content that our editorial skills can deliver. ■

"Our online trial opens up a parallel track of peer review for submitted papers for authors willing to go down that route."

RESEARCH HIGHLIGHTS

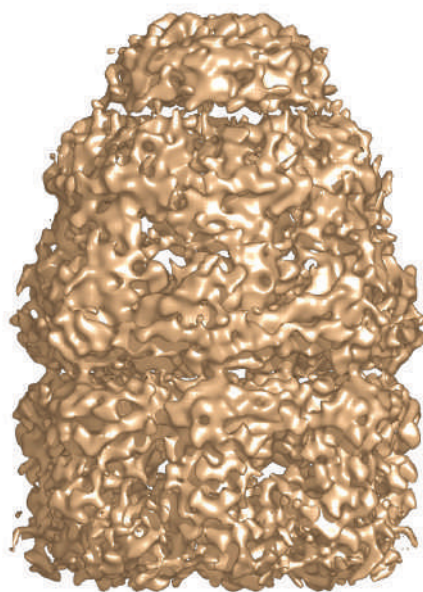
Size matters

Cell **125**, 903–914 (2006)

The molecular cage that provides a safe haven for folding proteins in *Escherichia coli* has more design features than was previously thought.

Manajit Hayer-Hartl and F. Ulrich Hartl of the Max Planck Institute of Biochemistry in Martinsried, Germany, and their colleagues tweaked various properties of the molecular cage — the chaperonin protein GroEL — and studied the effects on its role as a folding helper.

It was once thought that proteins held within the GroEL complex (pictured) were simply protected from interactions with other molecules, but proteins inside its cavity fold more quickly than they do in the open. The researchers show that both the size of the protein's caging cavity and its surface charge are important in optimizing folding speed.



RANSON ET AL. NATURE STRUCT. MOL. BIOL. **13**, 147–152 (2006)

CANCER BIOLOGY

Werner gene silenced

Proc. Natl Acad. Sci. USA **103**, 8822–8827 (2006)

The gene to blame for higher cancer prevalence in patients with Werner syndrome, an inherited disorder associated with premature ageing, also plays a role in cancers in patients that don't have the syndrome, according to new research.

Manel Esteller at the Spanish National Cancer Centre in Madrid and his colleagues looked at more than 600 human tumours of 12 different types. They found that the *WRN* gene, which is inactivated by mutations in patients with Werner syndrome, is silenced in tumours by the addition of methyl groups at its activation site.

WRN is involved in DNA repair, so its inactivation may enable tumours to acquire mutations that help them to survive, but it also makes them more susceptible to DNA-damaging chemotherapeutic drugs, such as irinotecan.

ASTRONOMY

Beyond expectations

Astron. Astrophys. doi:10.1051/0004-6361:20065476 (2006)

A piece of missing physics may explain why the sizes of planets found orbiting stars other than the Sun don't conform to the predictions of standard planetary evolution models, argue Tristan Guillot of the Cote d'Azur Observatory in France and his colleagues.

So far, the radii of ten massive planets have

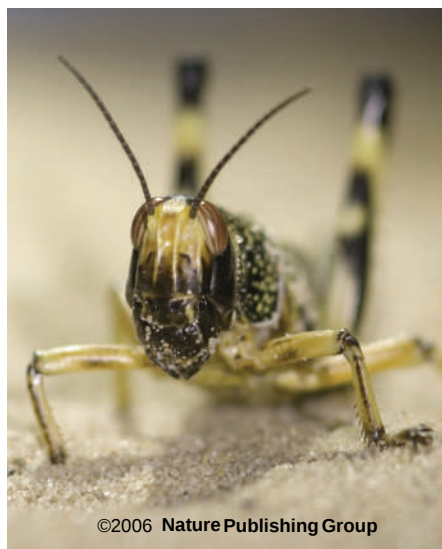
been measured by watching them 'transit' in front of their parent stars. Two of the 'hot Jupiters' were much larger than expected. The researchers say that a model which includes an additional energy source inside the planets to slow their contraction can explain the results. They also say that the model's predictions for the abundance of heavy elements in the planets supports the idea that this quantity is correlated with the metallicity of the host stars.

ANIMAL BEHAVIOUR

March of the locusts

Science **312**, 1402–1406 (2006)

"The locusts have no king, yet they advance in ranks," the Bible tells us. Indeed they do, and with terrible effect: marching bands of wingless, juvenile desert locusts (*Schistocerca gregaria*, pictured) can extend for kilometres,



and are the precursors to swarms that devastate crops. Understanding band formation is therefore critical to pest control.

Jerome Buhl and his co-workers at the University of Oxford, UK, have found that it occurs abruptly, in a kind of phase transition, once the locust density exceeds a critical threshold. Their lab observations reveal one transition from random to coordinated movement of individuals, then a second to unidirectional group motion. The behaviour is predicted by a simple model in which each locust tends to align itself with its neighbours.

SPACE SCIENCE

Space rock rendezvous

Science **312**, 1330–1353 (2006)

Data collected by Japan's Hayabusa probe during its visit to the asteroid Itokawa are presented in seven new papers.

Among the main findings reported by the international team of scientists is evidence that the asteroid is made of rubble. Hayabusa observed that Itokawa's gravity is much weaker than that of other similarly sized asteroids and revealed its surface to be extremely rocky. This suggests that Itokawa was formed by the reagglomeration of pieces of a bigger asteroid that had been broken up in a collision.

Hayabusa was meant to return fragments of the asteroid's surface to Earth, but there are doubts about whether the probe managed to collect any material during its landing in November 2005, and engine problems may thwart its journey home.

G. A. MILLER

BIOTECHNOLOGY

From curse to cure

Proc. Natl Acad. Sci. USA **103**, 8804–8809 (2006)
Scientists are edging closer to transforming tobacco (pictured right) from a cancer-causing to a cancer-curing plant.

Hilary Koprowski of Thomas Jefferson University in Philadelphia, Pennsylvania, and his colleagues infected tobacco plants with DNA that codes for antibodies against a tumour-signalling molecule. The molecule, a sugar marker known as Lewis Y antigen, is expressed by breast, colorectal and other cancers. As in previous studies, the plant-produced antibodies attacked tumour cells in culture and cut tumour progression in mice.

In this work, the researchers boosted their yield by fusing a genetic tag called KDEL to the code for the antibody. They also showed that the antibodies bound to the human Fc receptor — a key mediator of successful immunotherapy.

PHYSICS

Feel the force

Phys. Rev. Lett. **96**, 200402 (2006)

According to quantum theory, a vacuum is far from empty: it seethes with photons that pop into and out of existence. An object moving through these photons should feel a slight drag — one type of the ‘Casimir effect’ — that Woo-Joong Kim and colleagues at Dartmouth College in Hanover, New Hampshire, think should be possible to measure.

In their proposed experiment, one end of a reflective cavity is vibrated in the gigahertz range, sending vacuum photons — which cannot be detected directly — towards a cloud of sodium atoms held in the container. The collision would knock the atoms into a



lower energy level, making them emit a second burst of photons that would signal the existence of a dynamic Casimir effect.

CELL BIOLOGY

Sharing the inheritance

Proc. Natl Acad. Sci. USA **103**, 8209–8214 (2006)

Genetic material must be shared out before a cell divides, but what happens to proteins? Biologists from the University of Oxford, UK, propose that a mechanism related to the one that partitions DNA divvies up and positions proteins, too.

In the bacterium *Rhodobacter sphaeroides*, a cluster of proteins making up a chemotaxis signalling pathway (which helps to direct movement towards chemicals) is localized in the cytoplasm at the cell’s mid-point. Judith Armitage and her co-workers show that this cluster duplicates before cell division, with the two clusters positioning themselves at the future centres of the daughter cells.

The team shows that cluster division and positioning depend on the protein PpfA, which is homologous to ParA, a protein involved in DNA partitioning.

QUANTUM PHYSICS

Phased and confused

Phys. Rev. A **73**, 051601(R) (2006)

The supersolid — an elusive quantum state of matter — could be probed by looking at the noisy motion of atoms held in optical lattices, say researchers in the United States.

Optical lattices use lasers to arrange a gas of ultracold atoms into a periodic array, mimicking the structure of atoms in crystalline materials. It has been predicted that these trapped atoms could settle into a supersolid state — in which they flow, like a superfluid, without resistance. So far, the only evidence for supersolidity comes from difficult experiments on solid helium.

Vito Scarola of the University of Maryland in College Park and his colleagues show that the noise correlation function for atoms in an optical lattice could be instrumental in revealing the signature of supersolid order.

PHYSIOLOGY

Pee for protection

Nature Med. doi:10.1038/nm1407 (2006)

An antimicrobial compound produced by the body that protects our skin from bacterial attack also disinfects the urinary tract, a new study shows. This provides an additional line of defence in a system otherwise kept clean by its regular flush-throughs.

Annelie Brauner of the Karolinska University Hospital in Stockholm, Sweden, and her colleagues detected the antimicrobial peptide cathelicidin in human urine and showed that it is produced by the epithelial cells that line the urinary tract. They also showed that strains of *Escherichia coli* that had caused more invasive infections were more resistant to the cathelicidin LL-37.

JOURNAL CLUB

Srikanth Sastry
Jawaharlal Nehru Centre for
Advanced Scientific Research,
Bangalore, India

A statistical physicist finds a way into a tangled web of matter.

Taking a bold step can leave you in a mess. But here I discuss an instance where it seems to have been successful.

I’m interested in the states that materials can end up in, other than their equilibrium states.

Equilibrium phase diagrams of simple substances present a tidy enough picture, with well understood fluid and crystalline phases. But when nature chooses to cheat equilibrium thermodynamics, it can weave a tangled web.

The fascinating body of recent work on how rigidity arises in colloidal solutions — small particles suspended in a fluid — presents one such tangle. Diverse mechanisms have been proposed to explain the process, known as structural arrest, but a fully satisfactory picture is yet to emerge.

Researchers at Harvard University in Cambridge, Massachusetts, tackled this problem by measuring the structure and dynamics of a colloidal solution during phase separation. They analysed the results using ‘mode coupling theory’, which uses the static structure of a material to predict its dynamics (S. Manley *et al. Phys. Rev. Lett.* **95**, 238302; 2005). This was a bold move, as phase separation in a colloid induces its structure to evolve continually. Nevertheless, it seems to have worked.

They find that structural arrest

results from a ‘glass transition’ in the particle-rich regions of the colloid, meaning the particles’ dynamics simply become too sluggish, because of their high density, for them to continue to phase separate. Computer simulations back up this view (G. Foffi *et al. J. Chem. Phys.* **122**, 224903; 2005).

A few years ago, I analysed the thermodynamic stability boundaries of glasses, but it was not obvious how to probe them experimentally. Manley and colleagues’ work shows an elegant way to do so.

NEWS

EU science fund targets young guns

MUNICH

Ten years ago, few would have imagined a European Union-wide fund for basic research, distributed without restrictions on scientific creed or country. The idea that the notoriously bureaucratic European Commission would finance such a fund, with no strings attached, would have stretched credibility too far.

But two decisions last week moved the much-heralded European Research Council (ERC) closer to reality.

The European Council of Ministers has approved a budget of €7.5 billion (US\$9.7 billion) over seven years. And the fledgling ERC's scientific council has agreed on a first funding stream, open to any scientist embarking on an independent research career.

Soon after that, a second stream of funding will be opened to investigators in all disciplines. Individual investigators may apply for grants of up to €1 million, and there is just one criterion for selection: excellence in cutting-edge research.

"ERC grants will be like the National Institutes of Health's R01 investigator grants," says Fotis Kafatos, a biologist at Imperial College London and chairman of the ERC's scientific council. "Individuals can put in an application at any time."

European scientists are famously frustrated by the complicated requirements of European Union (EU) Framework research programmes. These generally require multidisciplinary, multinational research groups working towards a defined industrial or social goal. Many have advocated an ERC for decades, arguing that it would raise the quality of

research and provide a new source of funding for scientists in countries where national funds are paltry.

The question was always: who would pay? National research agencies have legal and psychological difficulties with putting part of their funds into an international pot. And the European Commission has no mandate to fund basic research, its aim being to promote industrial competitiveness and improve quality of life in Europe.

Mounting pressure from scientists eventually prompted a pragmatic solution, and the ERC is now part of the EU's next Framework Programme, FP7, which will run from 2007 to 2013.

On 30 May, the council of ministers approved a budget of €50.5 billion for that programme and agreed on how it would be broken down, including the ERC's share. The plan must be ratified on 15 June by the European parliament, but is expected to pass with no major changes.

Foot on the ladder

The ERC's funding is less than had been hoped for when the concept began to gain momentum in 2002. "But it is not crumbs," says Kafatos. The annual budget will start at €300 million, rising to €1.5 billion in the seventh and last year. By then, Kafatos hopes, the ERC will have proved its value to politicians, and future Framework programmes would receive at least that much money.

In the meantime, the ERC will open its

Starting Grant scheme in its first year, aimed at scientists who are within ten years of completing their PhDs and seeking their first post as an independent investigator.

"There are many countries where young scientists are kept dependent for far too long on their professors," says Kafatos. "This programme will allow them independence earlier."

Julia Fischer, an animal-behaviour specialist at the German Primate Centre in Göttingen, agrees. "There is a big difference between countries," she says. "In the United Kingdom it is easier to get on the career ladder than in Germany, for example." Fischer is on the board

of the Young Academy, which comprises 50 top young scientists in German-speaking countries.

One of the biggest concerns facing the ERC is the danger of oversubscription. There are well over a million full-time

researchers in EU countries.

"We have no idea how many scientists across the EU would in fact be eligible for the starting-investigators programme, or how many will apply," admits Helga Nowotny, a social scientist at the Vienna Centre for Urban Knowledge Management and vice-chair of the ERC's scientific council.

But, she says, the planned two-stage selection process will ease the pain of any oversubscription. In the first stage, applicants must summarize their projects and give information about the research team that would carry them out. The winners in this round will be invited within six weeks to submit a full application.

The council will establish at least 15 panels, organized by subject, to handle the final decisions. The panels will choose external referees to review the full proposals. This differs from the selection procedures of the Framework programmes, where the commission selects evaluators from a list of scientists who have volunteered for the job. "That is not a good system," says Nowotny.

The scientific council is now consulting organizations such as national granting agencies and academies to identify panel members. "Only those with strong scientific reputations will be selected to join panels," says Kafatos.

He doesn't think oversubscription will be a problem. "Scientists will exercise self-evaluation," he says. "Knowing that only frontier research is going to be funded, only those who have really excellent proposals will apply." ■

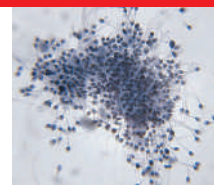
Alison Abbott

"In many countries, young scientists are kept dependent on their professors for far too long."



In supporting the European Research Council, the European Commission has bucked its reputation for bureaucracy.

T. MONASSE/AFP/GETTY



SENIOR SPERM ARE LESS FIT

Men, prepare to hear the ticking of your biological clock

www.nature.com/news

AP PHOTOS, BALDWIN



Hope out of tragedy: new identification techniques were developed after 9/11.

Disasters drive DNA forensics to reunite families

HELSINKI

Forensic tools developed because of the terrorist attacks of 11 September 2001 and the Indian Ocean tsunami of 26 December 2004 could soon find a happier use — helping reunite Jewish families separated by the Holocaust.

The DNA Shoah project, announced on 1 June at the Human Genome Organisation's meeting in Helsinki, Finland, is the brainchild of Syd Mandelbaum, whose parents survived the Holocaust, or Shoah as it is known in Hebrew.

Working with Michael Hammer, a geneticist at the University of Arizona in Tucson, Mandelbaum has launched a database of the DNA of survivors. The hope is to match sequences with Holocaust-era remains that have surfaced in land-development projects in Poland, Germany and elsewhere in Europe. Another aim is to reunite some of the families of around 10,000 Holocaust orphans sent abroad from mainland Europe to Israel, the United States, Britain and elsewhere between 1945 and 1950.

It is a difficult task. After more than 60 years in the ground, DNA will be degraded. And there are few reference sequences to compare it with: so many people were murdered that in many families all immediate relatives were killed. Matching distant relatives is statistically much harder.

Findings from more recent disasters could help. At the World Trade Center in New York, where more than 2,700 died in the 9/11 attacks, fire degraded DNA to the point that standard DNA fingerprinting could not give a reliable match (L. G. Biesecker *et al. Science* 310, 1122–1123; 2005). Investigators turned to mitochondrial DNA, which is more plentiful and robust than the commonly used nuclear DNA. They also adopted a method called SNP typing, which works with smaller pieces of DNA. Yet, individually, none of these methods was precise enough to identify most remains.

"It pushed the science to try more innovative techniques," says Richard Leary, managing director of the

company Forensic Pathways, based in Staffordshire, UK.

Gene Codes Forensics, a bioinformatics firm based in Ann Arbor, Michigan, developed software to integrate different kinds of DNA data and other forensic information. On the first day the system was used, 13 December 2001, it made more than 80 new matches, which helped identify 55 people who had been killed at the World Trade Center. So far, 1,598 victims have been identified by all means.

Identifying those killed by the Asian tsunami posed different problems. The wave swept away houses containing personal items, such as hairbrushes, that could provide DNA samples, and also killed many family members at once, says Kirsty Wright, a forensics expert at the Queensland Health and Scientific Services in Brisbane, Australia, who helped identify victims. As well as multiple DNA profiling methods, the teams needed software that could match victims with distant relatives.

Wright is now working on ways to

match the DNAs of dead people with each other, so that multiple victims in one family can be identified simultaneously.

The software from Gene Codes, which is among those that can match distant relatives, will handle the Shoah data. But time is running out. There are some 300,000 known Holocaust survivors, with an average age of 68. Mandelbaum and Hammer hope soon, with the help of the US Holocaust Memorial Museum in Washington DC, to start collecting DNA from cheek swabs.

For Howard Cash, president of Gene Codes, the idea that his firm's software could be used to bring good news is welcome. "Our team has dealt with such terrible circumstances," he says. "They would be overjoyed to reunite someone with a family member they thought was lost."

Leary says it highlights the potential of DNA forensics. "The issue for geneticists is to use this science in a very wide and imaginative way," he says. Claire Ainsworth

Insurers' disaster files suggest climate is culprit

MUNICH

Insurance companies, acutely aware of the dramatic increase in losses caused by natural disasters in recent decades, have been convinced that global warming is partly to blame. Now their data seem to be persuading scientists, too. At a recent meeting of climate and insurance experts, delegates reached a cautious consensus: climate change is helping to drive the upward trend in catastrophes.

The meeting, held near Munich on 25–26 May, was jointly organized by Munich Re, the world's largest reinsurance company, and the University of Colorado in Boulder. It brought together climate, atmosphere and weather researchers with economists and insurance experts to discuss what could be behind recent disaster losses, both economic and human. Insurers have been outspoken in blaming global warming. But scientists have tended to be more cautious, with many arguing that the rise could be primarily due to socioeconomic changes and natural climate variability.

Under discussion were data compiled by Munich Re, whose NatCatservice database, comprising 22,000 natural disasters dating

back to AD 79, is the largest of its kind. NatCatservice shows that the frequency of weather-related catastrophes has increased sixfold since the 1950s. The number of non-weather disasters, such as earthquakes, tsunamis and volcanic eruptions, has only marginally increased during the same period.

Delegates seem to have found the record persuasive. Their consensus statement, to be released on 8 June, says there is "evidence that changing patterns of extreme events are drivers for recent increases in global losses".

There was no agreement on how big a role global warming has played, however. "Because of issues related to data quality, it is still not possible to determine the portion of the increase in damages that might be attributed to climate change," the workshop concluded.

"Dissent over the issue is clearly waning," says Peter H  ppe, head of Munich Re's Geo Risks department, who co-chaired the workshop with Roger Pielke Jr, director of the University of Colorado's Center of Science and Technology Policy Research. "Climate change may not be the dominant factor, but it has become clear that a relevant portion of



damages can be attributed to global warming."

Previously sceptical, Pielke says that he is now convinced that at least some of the increased losses can be blamed on climate: "Clearly, since 1970 climate change has shaped the disaster loss record."

He adds a note of caution, however: "Disaster damage is not the place to look for early indications of climate change," he says. "Policy advocates should exercise caution in using

Politicians chastise Australia's science institute

SYDNEY

Australia's biggest research organization has faced tough questioning in the national parliament over its financial management after it had to reduce its estimates of future external earnings. But the organization says the fall is just a blip caused by legal tussles over its intellectual property and the fact that much of its income is now in the form of people and equipment rather than cash.

The Commonwealth Scientific and Industrial Research Organisation (CSIRO) has a turnover of nearly Aus\$1 billion (US\$750 million),

two-thirds of which comes from the government. Earnings from non-governmental sources, such as industry, are estimated to fall by nearly Aus\$184 million over the next four years — a humbling blow, critics say, to chief executive Geoff Garrett, who had pledged to boost such external revenue.

The opposition Labor party also attacked the agency, at a parliamentary hearing on 31 May, for frittering away funds on entertainment and the chief executive's home renovations. "The Australian community is rightly asking: what's going on with CSIRO?" says Jenny Macklin, Labor's

shadow minister for science. "It's time we got an answer."

This bad publicity for the agency follows earlier criticism of its controversial diet book

"'Hospitality' is not about executives going off to restaurants and drinking top-shelf wine."

and accusations that its climate scientists were prevented from speaking out about climate change (see *Nature* 438, 1060–1061; 2005 and *Nature* 439, 896–897; 2006).

But CSIRO management has hit back, arguing that there are

sound reasons for the revised forecasts. About Aus\$121 million of the shortfall is due to delays in earning money from the agency's patents, it says. Legal cases over patent infringement, including a battle with information-technology giants Microsoft, Dell and Hewlett-Packard over a patent on wireless technology, are partly to blame. "Until we know the outcomes of those legal cases, we have recalibrated so as not to take into account revenue from those sources," says CSIRO's chief finance officer, Mike Whelan. "When we know the outcomes, we will revise the estimates."



An ill wind: an April storm destroyed this house in Tennessee, and the weather seems set to worsen.

The group ranked countries according to four indicators — the number of casualties from extreme weather events, the number of casualties per 100,000 population, total economic damages, and economic damages relative to the country's gross domestic product. It combined those rankings to give a final "climate risk index". The ranking is topped by Somalia and other underdeveloped countries. But the United States and Japan, where natural disasters in 2004 caused considerable economic losses, also ranked highly.

The list has been criticized by experts, who say that data for a single year are largely influenced by random events and reveal little about how climate change is affecting different countries. But Sven Anemüller, co-author of the report and senior adviser for climate and development with Germanwatch, counters that the 2004 list is just the beginning of a longer-term analysis. "We're not saying that the countries listed this time are the ones that will lastingly be hit hardest by climate change," he says.

The group's approach, if continued, does make sense, agrees Höppe. Looking at the rankings over years or decades could eventually provide much-needed information as to which countries are most at risk from climate change, he says. ■

Quirin Schiermeier

disaster losses to justify climate mitigation, lest they go beyond what science can support."

But environmental groups are already using the rise in extreme weather events to help their campaigns about climate change. For example, Germanwatch, an environmental and developmental watchdog group, has used the Munich Re data to compile a ranking of the countries most badly affected by weather-related disasters in 2004.

Officials say the rest of the shortfall results from industry giving people or lab equipment, rather than cash, to collaborative research programmes. Whelan says this practice is increasing and isn't counted as revenue.

Of the Aus\$750,000 spent on renovating the chief executive's abode in Canberra, CSIRO officials say it was entirely reasonable to renovate the "run-down" building, which is heritage-listed, and to transform it from offices into a liveable residence.

They also took umbrage at accusations of corporate excess after it was revealed that the organization racked up a Aus\$1.4-million 'hospitality' bill in the past

financial year. According to Whelan, most of the money was spent on scientific conferences and meetings to plan research. "It's not about executives going off to restaurants and drinking top-shelf wine," says Whelan. "It involves bringing scientific collaborators together — and, yes, sometimes we give them a cup of tea."

Staff morale, already low according to annual staff surveys, may dip further with these recent developments, says Michael Borgas, president of CSIRO's staff association. CSIRO has undergone significant changes since Garrett took the helm in January 2001, with more priority-driven

research and pressure for commercial outcomes. The biggest concern, according to Borgas, is that the organization may be pushing ahead too quickly before having successfully persuaded staff of the need for change. Notably, no staff survey is being conducted this year.

Perhaps most galling for staff — who face more job cuts over the next few years — is the revelation that Garrett earns Aus\$478,980 a year, whereas senior researchers' wages typically fall below Aus\$100,000. And, according to Labor, more CSIRO senior executives than ever are earning over Aus\$300,000. ■
Carina Dennis

R. W. LOWE/USFWS



Do not disturb: Steller sea lions will enjoy a hassle-free summer

Sea-lion studies come to a halt after court judgement

A US federal court has shut down a research programme on the endangered Steller sea lion, ruling that studies — including the branding of pups — were inadequately monitored and could harm the animals.

The 26 May ruling scuppers the plans of groups of researchers who intended to study the sea lions (*Eumetopias jubatus*) this summer from Oregon to Alaska.

Last year, the Humane Society of the United States, an animal-welfare group based in Washington DC, sued the US National Marine Fisheries Service (NMFS), alleging that the agency failed to monitor and analyse studies worth more than \$120 million begun in 2001 (see *Nature* 436, 315; 2005). In her ruling, Judge Ellen Huvelle determined that the NMFS “arbitrarily and capriciously” failed to do so.

Sharon Young, the humane society’s marine issues field director, says the organization sued to force the NMFS to assess the environmental impact of the research. “This lawsuit was about conducting good science by coordinating and designing valid studies,” she says.

“We are disappointed in the ruling, but we respect it,” says NMFS spokeswoman Susan Buchanan. Agency officials say they have already begun the impact assessment, which will take at least another year.

This summer, biologists planned to ‘hot brand’ 2,900 Steller pups, identifying them for

later observation. Air surveys and collection of blood and other samples were planned. All fieldwork is now blocked.

Kate Wynne, a biologist at the University of Alaska, is disappointed with the halt to branding, which is, she says, “a valuable tool”.

Federal records that emerged during the lawsuit show that at least eight pups died after a branding expedition in 2001, which exceeds the allowed five accidental deaths. The NMFS attributed one death to that study. About 500 pups were born in the study area that year.

Roy Lowe, a refuge manager with the US Fish and Wildlife Service, says his team was worried about how many sea lions were dying. He later limited studies to alternate years, and this year turned down an NMFS request to brand and tag on the Oregon coast.

In April, the NMFS took the rare action of fining Randall Davis, a physiologist at Texas A&M University in Galveston, \$6,666 and denying him research permits until next year. He was accused of using an unapproved anaesthetic on Alaskan Stellers and of improperly conducted studies of pups in 2003 and 2004. Davis could not be reached for comment.

“Many of the scientists have themselves to blame for this,” says Young. “They need to look in the mirror, and hold themselves to the same standards they hold others to.” ■

Rex Dalton

ON THE RECORD

“I am the most loved woman in the world, the one with the smile full of mystery.”

The Mona Lisa introduces herself, thanks to a Japanese forensics expert who claims that he has recreated the sound of her voice.

“Like a 300-year-old Wal-Mart on the bottom of the ocean.”

Explorer Barry Clifford describes the huge number of artefacts recovered from the wreck of the pirate ship *Whydah* off Cape Cod, Massachusetts.

Sources: Associated Press, Los Angeles Times

SCORECARD

Asparagus
A UK botanist helps save wild asparagus by fertilizing one of the plant’s few remaining females with pollen from male plants 200 km away.

Publishing opportunities
Bird flu gets a journal all to itself, as Blackwell launches *Influenza and Other Respiratory Viruses*. For those tired of avian influenza, it will also, as advertised, cover a range of respiratory viruses.

Smallpox
World health officials postpone smallpox’s doom once more, as they fail to set a date for the destruction of the last stocks of the virus.

NUMBER CRUNCH

Drinking alcohol can make you do stupid things — and for some it can mean more than just a hangover. A few facts from a recent survey of drinkers attending hospital with injuries might give you pause as you reach for your glass.

3.8% of drinking-related injuries are intentionally self-inflicted.

14% of drinking-related injuries are intentionally inflicted by someone else.

32.4% of drinking-related injuries result from falling or tripping, the most common source of harm.

Source: Borges, G. et al. Bull. World Health Org. 84, 453–460 (2006)



CC STUDIO/SPL

It's all in the wrist: the finer points of experimental technique often never make it beyond scrawled notes or laboratory conversations.

Online methods share insider tricks

Replicating controversial lab results or tricky methods could become easier, thanks to a new breed of websites where scientists share and edit each other's laboratory techniques.

Laboratory protocols in biology and chemistry — the step-by-step guide to, say, separating proteins or splicing DNA fragments — are conventionally published in research papers or books of standard protocols. The instructions should allow another researcher to copy and confirm an experiment.

But scientists know that these recipes are seldom enough. Journals are cutting their methods sections to save money. And printed protocols lag behind rapidly evolving and increasingly sophisticated techniques, such as the nuclear transfer used in cloning.

Perhaps more importantly, it is the subtle variations — the deftness of touch, the type of mixing tube, and a dash of hocus-pocus — that distinguish a successful experiment from a flop. But such details often exist only as scrawled footnotes or collective laboratory wisdom. "The art of the science really is not present in many of these protocols," says geneticist Garry Nolan of Stanford University, California, who has put his protocols online. "They don't tell people what the voodoo is."

"Printed protocols lag behind rapidly evolving techniques such as those used in cloning."

The websites could help share the voodoo. They are loosely based on the online encyclopaedia Wikipedia, which lets users edit each other's entries. Unlike the protocols already available online, the idea is to create a repository of experiments and the tricks needed to do them, and allow users to add their own.

One burgeoning site, OpenWetWare, was set up just over a year ago by students at the Massachusetts Institute of Technology. The

Wikipedia-style site, featuring methods and other scientific resources, had around 30,000 users last month. One of the most popular protocols, used to measure the level of protein production in cells, now includes experimental data posted by users to let others

know what to expect.

"You can't find this information anywhere else," says one of OpenWetWare's founders, Sriram Kosuri, a graduate student in synthetic biology. The site is particularly popular among researchers in synthetic biology, who want to create standard tools for engineering biological systems.

Two other competing sites are starting up. One, from the Cold Spring Harbor Laboratory Press, launched last week, and one, from the Nature Publishing Group, publisher of *Nature*,

is due to launch in June. Both will feature commissioned protocols, which users will be able to comment on and add to. Unlike Wikipedia, comments will be screened before they are published and some of the material will be available only to subscribers.

Advocates say the sites have several advantages. They help busy lab heads deal with enquiries about their protocols. By removing some of the mystery from methods, they could help researchers iron out flaws, and perhaps verify controversial results. They might also raise the profile of methods, often glossed over as the means to the more exciting results.

On the flip side, a laboratory worker could end up drowning in information. And the sites will be successful only if enough scientists embrace them. Researchers in competitive fields might hold back methods that they think give them an edge. "A lot of molecular biologists are not very comfortable on the Internet to begin with," says systems biologist Pamela Silver of Harvard Medical School, who uses OpenWetWare.

And even the sites' supporters admit that a written protocol still cannot compare to learning on the job from a lab veteran. "The very best way," says John Inglis, executive director of the Cold Spring Harbor Laboratory Press, "is to sit beside someone who's doing it."

Helen Pearson


**MOM'S DIET CAN TINKER
WITH BABY'S GENES**

Get coverage from
the Human Genome
Organisation meeting at
www.nature.com/news

US scraps plan to license foreign scientists

Scientists in the United States have won a reprieve from a proposed security regulation that would have restricted some foreign researchers' access to laboratories.

On 31 May, the US Department of Commerce announced that it was withdrawing a proposed rule change that would have required scientists and students from "countries of concern" — including Pakistan, India, Russia and China — to be licensed before using certain pieces of laboratory equipment. The decision, which surprised many, was hailed as a victory by science advocates in Washington DC. They say that the department responded responsibly to hundreds of complaints from researchers.

The rule change was first suggested in 2004 by the department's independent inspector-general, as a modification to the existing 'deemed export regulations' on the export of sensitive technologies (see *Nature* **431**, 615; 2004). The changes would have required some

foreign researchers in the United States to obtain licences to use many pieces of equipment listed as 'controlled', because the devices could be used to strengthen a nation's military capabilities. An unclassified version of the list includes everything from chemical lasers to bacteria.

Science advocates warned the proposal would create a mountain of paperwork and restrict access to laboratories. "It would have changed the whole open-campus research environment," says Robert Hardy of the Council on Government Relations, a Washington-based association for research universities.

In a four-page notice in the *Federal Register*, the official noticeboard of the US government, the department said that an existing exemption for foreign scientists would continue for now. The decision was made partly in response to

complaints from universities, industry and even some government labs, according to David McCormick, undersecretary of commerce for industry and security. "We don't

want to create regulations that make the research enterprise less productive," he says.

McCormick says that the department is forming a 12-person committee, which he hopes will include experts in academia, industry and security, to examine the issue of laboratory security. The group, he says, will have a mandate to examine "fundamental issues to do with deemed export policy".

Tobin Smith, senior federal-relations officer at the Association of American Universities, says that he feels the group will be able to come to sensible conclusions about security issues. "I'm pretty optimistic," he says. ■

Geoff Brumfiel

**"We don't want to
create regulations
that make the
research enterprise
less productive."**

South Korea gives funding boost to stem-cell research

The South Korean government will give stem-cell researchers 430 billion won (US\$450 million) over the next decade, in a bid to recover from the scandal involving cloning researcher Woo Suk Hwang.

On 29 May, the Ministry of Science and Technology released a plan compiled by a group of 50 researchers and government officials as part of a strategic effort to make Korea one of the world's leading stem-cell research hubs. It aims to strengthen embryonic stem-cell research and boost adult stem-cell research and infrastructure.

The funding is less than scientists had asked for but still represents an increase on last year's budget. In addition, the government has provided more than 11 billion won for Hwang's group, money that had been approved before the Hwang scandal.

US halts work with South Korean stem-cell line

US officials have halted research involving a human embryonic stem-cell line from a South Korean team, after allegations that

unapproved cell lines had been shipped to researchers.

The cell line was one of those approved by the US government, because it was created before President Bush's 2001 ban on federal research on new stem-cell lines. The National Institutes of Health (NIH) has suspended its five-year contract with MizMedi Women's Hospital in Seoul for the line dubbed Miz-hES1.

About three-dozen US researchers at the NIH and universities had received the cell line in question, says John Burklow, an

NIH associate director of communications. His office is investigating a report that other cell lines were substituted in South Korea for some of those shipped as NIH-approved. MizMedi officials did not respond to enquiries.

Health agency puts focus on diseases of the poor

Global health campaigners scored a victory in Geneva on 27 May, as the World Health

Israeli cave reveals eight arthropod species

A unique and isolated ecosystem has been discovered in a subterranean lake in a cave 100 metres below a limestone quarry in central Israel.

Hanan Dimentman, a zoologist at the Hebrew University of Jerusalem, collected specimens of eight previously unknown arthropod species, including four kinds of crustacean, a springtail and a scorpion. DNA tests are being done to determine when the animals might have diverged from their marine and freshwater relatives.



The lake was discovered by a geography masters student, Israel Naaman, as he explored the groundwater beneath the quarry.

S. TIRAM

Assembly embraced a measure to tackle medical problems in poor countries.

The resolution directs the World Health Organization (WHO) to set a strategy to spur research on diseases that disproportionately affect the poor. Brazil and Kenya led the push for the measure, which endorses the work of a WHO commission on intellectual property and public health. A report published by the commission in April found that patent laws are one way — but not the only way — to encourage medical research (see *Nature* 441, 135; 2006).

The move is a step in the right direction, advocates say. Over the next year, the WHO's member states will try to work out a strategic plan, with a final report expected by May 2008.

Europe's space lab seeks date with Destiny

At long last, the science lab destined to be Europe's research outpost in space has completed the first leg of its journey into low Earth orbit.

The Columbus science module — a 7-metre-long capsule packed with racks of experiments — arrived at NASA's Kennedy



Hitching a lift: the European space lab Columbus went by plane to NASA's Kennedy Space Center.

Space Center in Florida on 30 May. The module was carried from its storage site in Bremen, Germany, in the belly of an enormous aircraft (pictured).

Columbus is expected to be launched aboard the space shuttle some time in the second half of 2007, after problems with the shuttle program pushed it back from 2005. The module will dock with the International Space Station, where its research facilities in biology, physiology and fluid physics will add to experiments already being done in orbit on the US science module, Destiny.

Transgenic drug gets the go-ahead in Europe

A drug made in the milk of genetically engineered goats may soon be used in a hospital near you.

Overtaking its previous decision not to approve the drug (see *Nature* 440, 13; 2006), the European Medicines Agency has approved ATryn for use in the European Union — the first time it has given the go-ahead for a drug from a transgenic animal source. The change in heart came after another detailed review of the data.

GTC Biotherapeutics, a Massachusetts-based company, engineered goats to express a human protein called antithrombin that stops the formation of blood clots. The protein, marketed as ATryn, will be given to antithrombin-deficient patients when they give birth or undergo surgery, where blood-thinning medicines could lead to runaway bleeding. The drug is also undergoing large-scale human trials in the United States.

Correction

The News story "Old tools shed light on hobbit origins" (*Nature* 441, 559; 2006) said that the hominids *Homo floresiensis* were ancestors of *Homo erectus*; it should have said "descendants".

K. SCHNEIDER/SPL



THE GENE WEAVERS

Viruses are often thought of as simple creatures. But their staggering diversity and genetic promiscuity could make them the most creative force in evolution, says **Garry Hamilton**.

When high-school students Joe Gross and Jake Falbo walked into Graham Hatfull's lab at the University of Pittsburgh five years ago, they had no idea what they were getting into. Hoping for a little hands-on scientific experience, the two were instead invited to participate in a mission that would have excited the likes of Linnaeus or Darwin. "I said to them, 'Sure you can do a project'," recalls Hatfull. "Why don't you go out and discover some new viruses?"

It was no joke. Gross and Falbo are among the co-authors of a now heavily cited *Cell* paper for their separate identification of two previously unknown viruses called bacteriophages, 'phages' for short, which infect bacteria¹. But unlike Linnaeus and Darwin, Hatfull's team doesn't have to embark on a far-flung expedition to collect and classify new creatures — they, and other virologists like them, come across them wherever they go. "All you need to do," says Hatfull, "is go and look under a rosebush and you'll find a phage that's unlike anything anyone's ever seen before."

In total, Hatfull and his squadron of phage-hunters have now isolated and sequenced

more than 40 phages, from settings as varied as the grounds of a tuberculosis clinic in India, the monkey house at New York's Bronx Zoo and, yes, even beneath a rosebush in Latrobe, 55 kilometres southeast of their Pittsburgh base. They have discovered that every phage they come across is almost completely different from the next. Findings such as these have prompted a huge change in how biologists see all viruses: the spurt has fuelled the suggestion that viruses have a large part to play in the evolution of life on Earth.

In from the cold

Viruses have long been viewed as nature's outsiders. As parasites that depend on a host cell for survival, they don't seem to have fully earned the stripes of a living organism. But they are actually far more abundant, diverse and complex than once thought. Recent calculations suggest there may be more undiscovered genes in the viral world — most belonging to phages — than in all other life forms combined. And this vast presence and diversity has profound effects on the rest of life. By shuttling genes into and out of their hosts, viruses seem to be a major driving force

in the evolution of higher organisms. Even within our own genome, genes that came from viruses are hard at work. This in turn has led to the realization that viruses probably play a major role in the ecological, biochemical and evolutionary processes that underlie the entire natural world. Rather than being the outsiders, viruses have emerged as perhaps life's most ubiquitous presence.

"I don't think you can look at any system now and leave viruses out of the equation," says Nicholas Mann, a microbiologist at the University of Warwick, UK.

Hints of the vast, unexplored viral world first emerged during the 1980s. Researchers at the University of Bergen in Norway, with the help of a new electron-microscopy technique, showed that viral concentrations in some aquatic habitats were up to 10 million times greater than previous estimates². Viral particles per millilitre ranged from 60,000, deep beneath the Barents Sea, to 254 million, from the surface waters of Germany's Lake Plüsee.

Since then researchers have discovered huge numbers of viruses wherever they have looked, from 2,000 metres below the surface of Earth to the sands of the Sahara Desert, from acidic hot

HIDDEN TALENTS

The discovery that we share the planet with countless viruses has led researchers to explore their strange talents.

What happens when this viral world throws its weight around, for example? By killing bacteria en masse, viruses are thought to play an important role not only in the control of microbial populations, but also in global geochemical processes such as the carbon cycle. Other research has shown that when viruses prey on blooms of marine algae, the volume of dimethyl sulphide released is sufficient to influence the formation of clouds¹².

And viruses seem to have other mysterious abilities. Many viruses carry and use a long list of genes formerly thought to be useful only in cells, such as those needed for

photosynthesis. They probably provide viruses with an evolutionary advantage while they are infecting cells, but it's not always clear how.

At the University of Nebraska in Lincoln, plant virologist James Van Etten has cultured an algae-infecting virus that carries a gene for hyaluronic acid, a molecule used by vertebrates as a joint lubricant, as well as viruses that make chitin, the protein used as a building block in insect exoskeletons¹³. "We know they use these genes," says Van Etten. "But we still have no idea what advantage they gain from having them."

Meanwhile, other recent finds have shattered the notion that all viruses are lean and simple. Most notable is the recently sequenced



M. HILL/ALAMY

Viruses thrive everywhere, from polar ice caps to hot springs.

mimivirus, an amoeba-infecting virus whose genome contains 1.2 million base pairs and more than 900 genes¹⁴. The mimivirus's DNA load, more than double that found in the smallest known

bacterial genomes, contains numerous genes long thought to be the exclusive property of cells, including genes for repairing DNA and manufacturing proteins. **G.H.**

spring to polar lakes. In total there are now thought to be some 10^{31} viral particles on the planet — an astronomical figure that one researcher recently described as 250 million light years' worth of viral genes laid end to end.

Then, in the mid-1990s, researchers started to grasp the dazzling variety that exists within this viral multitude. Biologists discovered that the known and cultured viruses represented just a fraction of their respective groups. From one cubic metre of seawater to the next, there was more genetic diversity in viruses than found in any other known group of organisms.

Constant reinvention

More recently, researchers led by microbiologist Forest Rohwer at San Diego State University, California, have developed a technique for extracting and sequencing unknown viral DNA en masse. In a series of studies that examined samples from seawater, marine sediment and human faeces, the researchers discovered that most viral groups had previously been completely undetected³. A kilogram of marine muck was found to contain up to a million different viral genotypes. In the human gut alone there may be as many as 1,200 distinct viruses. "Viral diversity is just way different from anything else," says Rohwer. "Every time we sequence it, most of everything we run into is unknown."

Hatfull and his colleague Roger Hendrix have made similar discoveries, with help from a team of high-school students and undergraduates. In the 40 or so phages they have found, roughly half of each genome consists of genes that have never been seen before — not in any of the cellular organisms that have been sequenced to date, not even in any other virus¹. And a similarly high percentage of

unique genes has been found in most other sequenced viral genomes — there are some 450 in the largest virus, the mimivirus (see 'Hidden talents'). This has led to the conclusion that the bulk of nature's genetic information resides in the genomes of viruses. "You could argue that this is nature's greatest genetic experiment," says Hatfull.

And viruses seem to be truly talented experimentalists. Although phages have long been known to cut and paste their genes together from different sources, Hatfull and Hendrix have shown just how rampant this shuffling is. By making a comparison of 14 different phages, the researchers detected spliced-in sequences that were no more than a single gene in length. Other segments were flanked by fragments of genes, as if they had been arbitrarily ripped from one virus and inserted into another. This suggests that, unlike cells, phages can combine bits of DNA even when there is no similarity between the sequences of the different DNA pieces.

Hatfull and Hendrix now argue — and many others concur — that viruses are continually and randomly recombining with whatever DNA they might encounter while infecting a cell, be it phage or host DNA. Thus cells are stewing pots where viruses are continually reinventing themselves thanks to a blind but highly creative process that is not only spitting out novel combinations of genes, but is also generating brand new genes possibly never before seen in nature. Success relies on the huge number of new viruses being created in the world — estimated to be some 10^{24}

per second. "Almost all of this would be junk, useless monsters," says Hendrix. "But it's happening often enough that the few that survive are still a significant number. It's darwinian evolution on a grand scale."

DNA dispersal

Hendrix also points out that the common structure of most phages — a capsule full of DNA and a hollow tail used for injecting that DNA into a cell — would be ideal for facilitating the creative process by requiring excess DNA to act as stuffing that enables the head to maintain its shape. Free from the pressures of selection, such non-essential DNA would have an extended opportunity to evolve into something useful.

The emerging picture suggests viral DNA can spread far and fast. Two years ago Rohwer and his colleagues reported that they had found the same sequence of viral DNA in 49 of 66 widely varied habitats — including the rumen of a cow from Idaho, a hot spring in California, the Antarctic Ocean and mucus from coral growing in the Caribbean Sea⁴. Analysis of 18 of these samples showed that the 533-base-pair sequence differed by no more than three nucleotides. Based on known phage mutation rates, this piece of viral DNA seems to have dispersed sometime in the past 2,000 years⁵.

It is possible that the same host resides in all of these habitats, or that a single virus can infect a wide range of hosts. But Rohwer and others, who have now found similar results for several sequences, believe it is more likely that fragments of viral DNA are being passed

"A viral group is like a super-organism — genetic information is shared between all these different viruses."
— Curtis Suttle

between viruses that share the same host.

"When you look at a group of viruses, such as the algal viruses, there seems to be a very, very small core of conserved genes," says Curtis Suttle, a microbiologist at the University of British Columbia in Vancouver, Canada. "The rest is almost like a super-organism — a massive pool of genetic information that's being shared among all these different viruses."

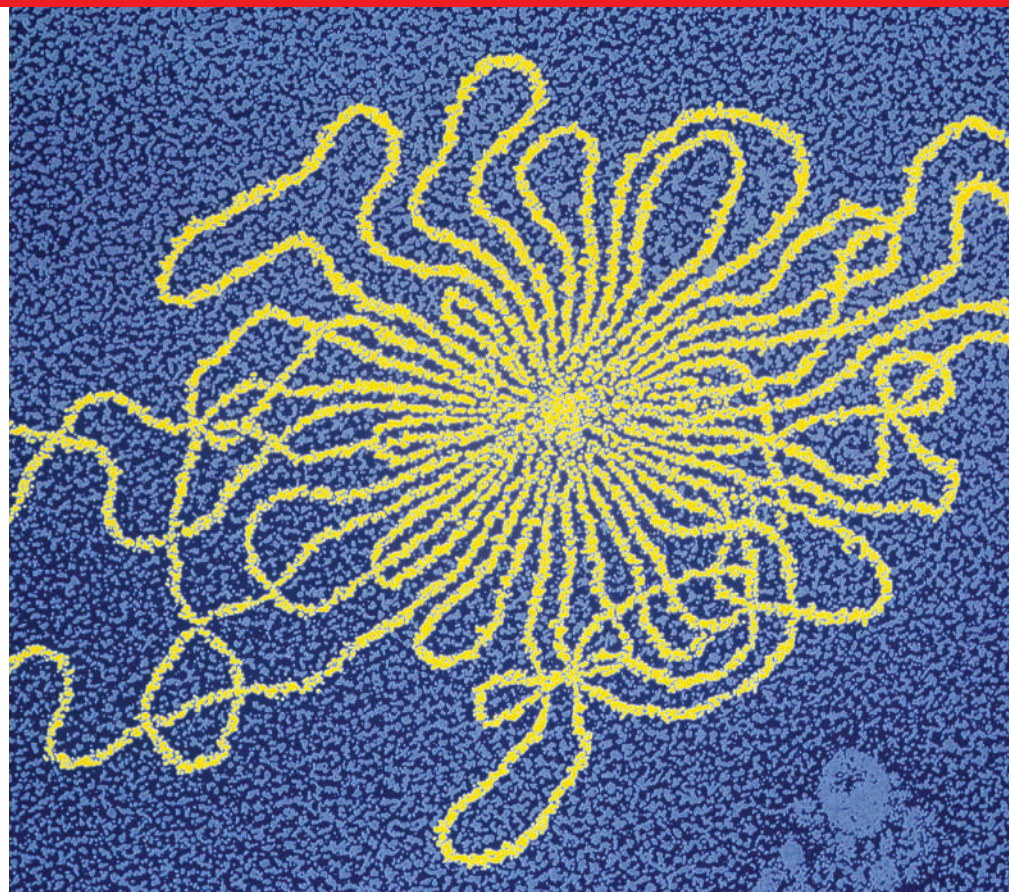
A big question now is the degree to which this super-organism extends its tentacles into host genomes. It has long been known that bacteria use genes acquired from prophages — phages that insert their DNA temporarily or even permanently into the DNA of their host — to gain competitive advantage and exploit new environments. Indeed, work reaching back decades has shown that prophage genes carried by bacteria are responsible for producing the primary toxins associated with diseases such as diphtheria, scarlet fever, food poisoning, botulism and cholera.

Behind bacteria

More recently, whole-genome sequencing has shown that most bacteria harbour an average of two or three prophages. Other studies have shown that phage genes make up the major differences between closely related strains of bacteria. In Japan, for example, researchers carried out whole-genome comparisons between a harmless laboratory strain of *Escherichia coli* and O157:H7, the strain of *E. coli* that has emerged in recent decades as a global health threat. They found that most of what separates the pathogenic strain from its close and harmless relative — almost a million base pairs of DNA — comes from 24 different prophage or prophage-like genetic segments⁶.

A more recent genome comparison between pathogenic and non-pathogenic isolates of *Neisseria meningitidis*, a normally harmless nasal-dwelling bacterium that is also known to cause meningitis, found only one noticeable difference between the two groups: a cluster of prophage genes that turned up in 29 of the disease-causing isolates but only 2 of the 20 benign bacteria⁷. This adds to a list of similar studies, such as one involving *Streptococcus pyogenes*, a resident of the skin and mouth that can cause a range of ailments including scarlet fever, rheumatic fever and toxic shock syndrome. In the study, researchers demonstrated how strains associated with three different disease processes each have their own distinct cocktail of phage-encoded toxins⁸.

But pathogenicity is just one trait that can affect bacterial evolution; scientists are beginning to find evidence that viral influence on the evolution of life may be a more general phenomenon. The bacterium *Pseudomonas aeruginosa*, for example, kills its competitors



Genetic cradle: this bacteriophage's DNA (yellow) could be a breeding ground for new genes.

with compounds produced by two modified phage genes⁹.

Scientists don't actually have to go as far as deadly bacteria, or even the nearest rosebush, to get new viral DNA. Human beings are full of it. Retroviruses are a type of virus that specialize in attacking animal cells. And roughly 8% of our genome consists of DNA copies of these RNA-based viruses that have incorporated themselves into our genetic make-up. One example is a protein used by viruses when they attach themselves to a host cell at the start of infection: it has recently been demonstrated to play an active role in binding together cells during development of the placenta¹⁰.

"Viral diversity is just way different from anything else. Every time we sequence it, we run into the unknown."
—Forest Rohwer

Evolutionary thoughts

Phage DNA has also made it into the human genome, despite the fact that phages target bacteria. The source was our mitochondria, the energy-producing capsules found in our cells and thought to be the descendants of a formerly free-living bacterium. During evolution, phage genes from the mitochondrial genome have transferred into our main genome, housed in the cell's nucleus. There, they help copy and express the few bacterial genes still present in the mitochondrial genome. Recently, this same phage DNA has been spotted in various modern bacteria belonging to the group from which mitochon-

dria are thought to have descended¹¹. That supports the idea that the gene made its way from virus to bacterium to cell nucleus, where it now plays a key role in the molecular circuitry that drives all multicelled organisms.

Some researchers now believe that viruses have been instrumental in assembling the various molecular components that define the cell types associated with life's three domains — bacteria, archaea and eukaryotes. Luis Villarreal, director of the Center for Virus Research at the University of California, Irvine, even argues that a large portion of what distinguishes humans from chimps is viral DNA.

"I think it's become apparent that viruses are involved everywhere," says Villarreal. "I would argue they are the most creative genetic entities that we know."

Garry Hamilton is a science writer based in Seattle, Washington.

1. Pedulla, M. L. et al. *Cell* **118**, 171–182 (2003).
2. Bergh, Ø. et al. *Nature* **340**, 467–468 (1989).
3. Breitbart, M. et al. *Proc. Biol. Sci.* **271**, 565–574 (2004).
4. Breitbart, M. et al. *FEMS Microbiol. Lett.* **236**, 249–256 (2004).
5. Breitbart, M. & Rohwer, F. *Trends Microbiol.* **13**, 278–284 (2005).
6. Hayashi, T. et al. *DNA Res.* **8**, 11–22 (2001).
7. Bille, E. et al. *J. Exp. Med.* **201**, 1905–1913 (2005).
8. Beres, S. B. et al. *Proc. Natl Acad. Sci. USA* **99**, 10078–10083 (2002).
9. Nakayama, K. et al. *Mol. Microbiol.* **38**, 213–231 (2000).
10. Mallet, F. *Proc. Natl Acad. Sci. USA* **101**, 1731–1736 (2004).
11. Filee, J. & Forterre, P. *Trends Microbiol.* **13**, 510–513 (2005).
12. Fuhrman, J. A. *Nature* **399**, 541–548 (1999).
13. De Angelis, P. L., Jing, W., Graves, M. V., Burbank, D. E. & Van Etten, J. L. *Science* **278**, 1800–1803 (1997).
14. Raoult, D. et al. *Science* **306**, 1344–1350 (2005).

OMIKRON/SPL



CULTURE CRASH

Some economists had hoped that physicists might shake up the rigid theories typical of mainstream economics. But so far, they're unimpressed by physicists' handling of the markets. **Philip Ball** reports.

For the past two decades, some physicists have been trying to apply their ideas and tools to an area that seems a long way from traditional physics. They are exploring the notion that there might be a kind of physics of the economy — an 'econophysics', as it has been dubbed¹. Last year, some of these econophysicists even went as far as to suggest that economics might be "the next physical science"².

But now this unlikely marriage is showing signs of turning sour. Even those economists who at first welcomed econophysics are starting to wonder whether it is ever going to deliver on its initial promise. Early successes in modelling financial markets have not led to insights elsewhere, some complain. Matters came to a head at the Econophysics Colloquium, held at the Australian National University in Canberra last November. A group of economists attending the meeting were so dismayed with what they saw many physicists doing that they penned a forthcoming paper entitled 'Worrying trends in econophysics'³.

In their critique, economist Paul Ormerod of the London-based consultancy Volterra and his co-authors accuse econophysicists of a litany of sins: applying inappropriate assumptions to economic systems, failing to do their homework properly, getting fixated on a small corner of the subject, and being sloppy with

their statistics. At face value it is a damning indictment, and raises the question of whether econophysics will ever make a genuine contribution to economic theory, or whether it is doomed to remain a fringe interest.

Claim to blame

Some econophysicists admit that there are problems. "Econophysics is a field with very uneven quality," says Doyne Farmer, a physicist at the Santa Fe Institute in New Mexico, who made pioneering contributions to the study of chaos before moving into economics. Yi-Cheng Zhang of the University of Fribourg in Switzerland is even more ready with a *mea culpa*. "My economist friends are right. The literature is often littered with garbage. We can find gauge-field theories of finance, quantum options and so on. In short, anything goes."

But others reject the accusations. In response to the Canberra critique, Joe McCauley, a physicist at the University of Houston, Texas, who now works mostly on economic problems, says, "Would one write an essay called 'Worrying trends in physics' simply because a few minor researchers put out bad papers? Bad papers, even wrong papers, appear in every issue of every scientific journal."

It is tempting to interpret this as a mere academic turf war. But Ormerod and colleagues are among the few people in economics who have taken econophysics seriously. Most economists don't know the discipline exists — and if they did, they would probably heap derision on it.

The idea that physics might have something useful to contribute to economics arises because both fields are concerned with systems of many interacting components that obey specific rules. Statistical physics describes the behaviour of bulk matter based on the forces acting between atoms and molecules. Economics studies the interactions between economic



"Econophysics is a field with very uneven quality." — Doyne Farmer

CHINA/GETTY



M. OESER/AFP/GETTY

'agents' — market traders, say, or businesses.

Arguably, deriving microeconomic principles from the behaviour of individual agents should pose similar problems to deriving thermodynamic laws from interatomic forces. The rules dictating how interactions play out between economic agents are admittedly more complex than the forces between atoms, but in conventional economics the rules have always been grossly simplified to make the models workable.

For example, the core theory of mainstream economics, the neoclassical model, argues that agents always act with perfect rationality to maximize their 'utility' (for example, profit), based on complete information about the state of the market as a whole. In this picture, an economic market quickly reaches an equilibrium state, in which commodities find the price that perfectly balances supply and demand.

Model world

Economists recognize that real human agents do not always act in such a coldly rational way, and that they generally have to manage with incomplete information. But although Nobel prizes in economics were awarded in 2001 and 2002 for work that recognizes these limitations, neoclassical theories — and particularly the idea of equilibrium — remain central to mainstream economics.

Ormerod and his colleagues, and other physics-friendly economists, had hoped that econophysics would help them create a new economics that is free from some of the dogmatic assumptions characterizing the mainstream discipline today. "Economics desperately needs econophysics," claims Ormerod's co-author Steve Keen, an econo-

mist at the University of Western Sydney in Australia. Keen had hoped in particular that econophysics might break his fellow economists' misperceived obsession with equilibrium. "Equilibrium thinking still has them in its unshakeable thrall," he says.

A glance at almost any plot of commodity prices over time belies the idea of market equilibrium: the values fluctuate wildly. But in neoclassical theory, these fluctuations are regarded as background 'noise' caused by unpredictable 'shocks' from outside the economic system, to which the market constantly and quickly adjusts. An attempt to explain such fluctuations using statistical physics was made as early as 1900 by Louis Bachelier; he proposed an explanation that introduced the theory of random walks, which was later developed independently by Einstein to explain brownian motion.

Bachelier's theory was deemed too strange to be taken seriously by economists. But the issue



"Physicists suffer from a belief that there must be universal rules."
— Paul Ormerod

Econophysicists have focused on the financial market because it offers a lot of good-quality data.

was revisited in the early 1960s, when mathematician Benoit Mandelbrot showed that fluctuations in cotton prices have a statistical distribution that differs from that expected of a typical gaussian process — where each event happens randomly and independently of all others. There were more large fluctuations than a gaussian distribution predicts⁴. This has significant implications for economic theories that assume market 'noise' to be gaussian — but more importantly, it suggests that big fluctuations, perhaps even market crashes, are not rare anomalies but intrinsic to normal market behaviour. Mandelbrot's 1963 paper on price fluctuations is now regarded as one of the key precursors to modern econophysics.

Out of equilibrium

But it wasn't until the early 1980s, when an unusual mix of researchers got together at the Santa Fe Institute, that economists showed much interest in scientific ideas related to complex systems. They were helped by advances in computing. "Once we got desktop computers, we could model systems of many agents and allow them rules of behaviour and see how they evolved," says economist Brian Arthur, who worked at Santa Fe alongside physicists and evolutionary biologists to develop non-traditional approaches to economics⁵. Such computer simulations of the economy led to models of 'interacting agents'⁶ that were influenced as much by work on cognition and evolutionary biology as by physics.

In these models, researchers could give the interacting agents any decision-making strategy they desired, and therefore study markets with different underlying behaviours. "What we found was quite surprising," recalls Arthur. "Under some restrictive conditions, you get market equilibrium, but under other conditions you get much more complicated outcomes." It seemed there was no good reason to believe that microeconomics always operates at equilibrium. "The economy is out of kilter most of the time," says Arthur. That, he says, accounts for one of the virtues of econophysics. "The core of economic theory is still built around equilibrium models, but most models in econophysics are non-equilibrium ones."

But those economists who have adopted new approaches such as agent-based modelling have become increasingly frustrated with the intransigence of mainstream economics. Some have even resorted to starting their own publication, the *Journal of Economic Interaction and Coordination*, the first issue of which appeared online in May. Zhang says that three of the four authors of the Canberra critique are victims of the intellectual exclusion imposed by mainstream economists. "That's why they had such high hopes when physicists offered what

IS ECONOPHYSICS REALLY NEW?

Although the idea of economists borrowing concepts from physics might seem unlikely, it has been a feature of economic theory ever since its inception. Adam Smith (pictured) wrote his *Wealth of Nations*, which laid the foundations of economic thought in 1776, in an intellectual climate that was infused with newtonian ideas about causative forces. Indeed, some of his contemporaries compared the circulation of commerce to

the circulation of the planets.

Pierre-Simon Laplace and the Belgian astronomer Adolphe Quetelet helped to establish the idea that there are natural laws, akin to Newton's laws of motion, that govern human social systems such as the economy. And nineteenth-century economic thinkers such as John Stuart Mill and Karl Marx



frequently alluded to scientific ideas and analogies.

Microeconomic theory, which strives to understand economic phenomena by building them up from the behaviour of individual 'agents' in the economy, was established in the late nineteenth century — just as James Clerk Maxwell and Ludwig Boltzmann were laying

the foundations of statistical physics. The early microeconomists Francis Edgeworth and Alfred Marshall drew on some of the ideas of these physicists, in particular the notion that the economy achieves an equilibrium state like that described for gases by Maxwell and Boltzmann. Edgeworth and Marshall's microeconomics led to the neoclassical theory of economics that defines the mainstream today. **P.B.**

CLASSIC IMAGE/ALAMY

seemed to be an alternative."

So why have some of these physics-friendly economists become fed up? Although Ormerod and colleagues are highly critical of mainstream economic theory, they point out that "economics is not at all an empty box." The Canberra critique accuses econophysicists of ignoring the existing literature — a charge also levelled at physicists when they began to dabble seriously in biology.

Zhang agrees this is a fair complaint. "Most econophysicists don't care about the underlying idea, history and cultural background to the problem. All they want is to get a lot of data, crunch the numbers better than anybody else, and publish a lot of papers." But he admits that doing the job properly is very daunting: "Compared to anything I've done in physics, the requisite reading is enormous. I get a pile of books delivered every month, and I'm just overwhelmed."

Another criticism raised by Ormerod and colleagues is that econophysics has too narrow a focus. Physicists have made important contributions to finance and industrial economics, Ormerod acknowledges, but he wants them to bring an open mind to other areas of the economy. "I think they fail to appreciate that financial markets as such are not terribly important in economic theory," he says. "They are just a special case of a more general theory of markets."

Why are physicists so finance-centred? Ormerod suspects that they feel more comfortable with the generally high-quality and long-term data from financial markets. Most other economic data exist in small, noisy data sets. Some econophysicists defend the focus on finance precisely for these reasons: "Finance data are strong enough to falsify specific classes of models," says McCauley. Farmer adds that physics has often been successful by tackling problems that could be solved with the data and the tools to hand.

Perhaps more damning, Ormerod and colleagues worry that some econophysicists have imported concepts from physics that just don't belong in economics. For one thing, they say, some have tried to apply conservation laws to quantities, such as money, that the economy

doesn't actually conserve. "Conservation of money is too far from the truth in a credit economy like ours," McCauley agrees. But he says that such mistakes are rare.

More broadly, claim the Canberra authors, physicists fail to appreciate how diverse and changeable the economy is in practice. "Physicists suffer from a belief that there must be universal rules," says Ormerod. "This is not a handicap in the physical world, but it is in economics, where the behaviour of agents may differ across time and place." But Farmer thinks that looking for generalities can still be useful: "The typical view among social scientists is that one should focus on documenting and explaining differences. Physicists have jumped in and said, 'Before we do that, let's spend some energy on first trying to understand if there is any way in which everything is the same'."

Out on a limb

Economist John Sutton at the London School of Economics agrees that physicists can help to identify relationships within economics that are driven by, say, elementary statistics, irrespective of any assumptions about agent behaviour. He says he welcomes outsiders to economics "who may look in different places for interesting relationships".

Sutton is, however, a rare example of a mainstream economist who is aware of the econophysics literature. Arthur believes that some shortcomings of econophysicists may be caused by their enforced isolation: "They tried to publish in mainstream economics journals, but were rebutted. It's a shame and a scandal that the journals haven't opened their pages to them." Consequently, econophysicists set up their own journals. This means they don't get the contact with economists that would have "rounded off their rough corners", Arthur says.

Perhaps not surprisingly, some econophysicists reject the Canberra critique outright. "The problem is not with econophysics," McCauley says. "The problem is within the economics profession." He argues that the neoclassical model, which is still taught in all economics

classes, was falsified in the 1970s by M. F. M. Osborne of the Naval Research Laboratory in Washington DC, who he regards as the first econophysicist (see 'Is econophysics really new?'). "Economists are overwhelmingly unaware of this contribution. Worse, the majority have no interest in anything that physicists write."

Arthur counters that economics has a long history of assimilating ideas from outside the field. But he explains that while economics tolerates some 'unorthodoxy', it also isolates it from the core theory that is taught to students and practised by academics. There is an attitude of "fine, but not in my journal," he says.

Despite their concerns, the authors of the Canberra critique still hold out hope for econophysics. "It's just that we don't want it going off the rails," says Ormerod. Yet for McCauley, the goal of econophysics is not to batter its way into the citadel, but to raze and rebuild it. "Econophysics will displace economics in both the universities and boardrooms," he says, "simply because what is taught in economics classes doesn't work."

Still, if econophysics is to survive and prosper, it needs support from somewhere, says Farmer. "It is clear that this is not going to come from economists." Turning to the physics community for help has not been the answer either. Rosario Mantegna, an econophysicist at the University of Palermo in Sicily, says that in academia physicists cannot pursue econophysics as their main research. "There is no one in the United States who has gotten tenure based on work in econophysics," Farmer agrees. "This is a real pity for physics."

Philip Ball is a consultant editor for Nature.

"Econophysics will displace economics in both the universities and boardrooms." — Joe McCauley

1. Mantegna, R. N. & Stanley, H. E. *An Introduction to Econophysics* (Cambridge Univ. Press, 2000).
2. Farmer, J. D., Shubik, M. & Smith, E. *Phys. Today* September 37–42 (2005).
3. Gallegatti, M. et al. *Physica A* (in the press).
4. Mandelbrot, B. J. *Business* **36**, 394–419 (1963).
5. Anderson, P. W., Arrow, K. J. & Pines, D. (eds) *The Economy as an Evolving Complex System* (Addison-Wesley, Reading, Massachusetts, 1988).
6. Tesfatsion, L. & Judd, K. (eds) *Handbook of Computational Economics Vol. 2* (Elsevier, Amsterdam, 2006).

BUSINESS

Biochemist strikes gold

Merck is tapping into academic research in its hunt for drug candidates. Meredith Wadman reports on the company's latest deal with Harvard.

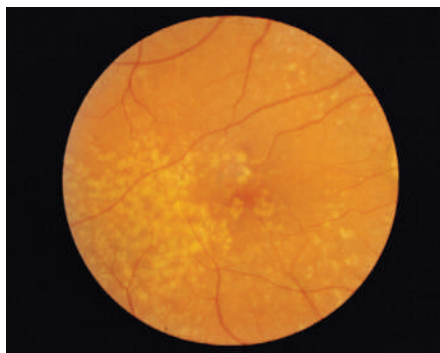
It's the stuff of academics' daydreams: do the science that opens a route to drugs for a common, debilitating and, so far, untreatable condition. Then have a megacompany swoop in and pay you handsomely for the privilege of — just maybe — turning your labours into a cure.

But in this case, it's no daydream. Late last month, it was announced that Harvard University and one of its biological chemists, Robert Rando, had signed a \$3-million-plus deal with Merck, aimed at developing drugs to battle a leading cause of blindness.

Under the agreement, the New Jersey drug company gets exclusive licenses to a family of candidate small-molecule drugs that Rando, a specialist in the biology of vision, has discovered and patented. The company hopes to find a drug that slows or prevents age-related macular degeneration (AMD) — a disease of the retina that is currently destroying the vision of nearly two million people in the United States alone¹.

The deal includes undisclosed future milestone and royalty payments from Merck if the company succeeds in developing a drug. "This is perfect," says Rando, "because I absolutely can't do what they are going to do. They are hopefully going to take this basic work and turn it into a drug."

The decision to strike such a deal reflects Merck's determination "to find innovation wherever it's occurring", says Robert Gould, the company's vice president of licensing and external research. In this case, he adds, it was hap-



Foresight: Merck strikes a deal that may help find a cure for a retina disease that affects millions.

pening "across the street at Harvard". Gould is based at Merck's two-year-old, 11-storey research laboratory in downtown Boston, right opposite the Harvard Medical School complex (see page 780).

Under Harvard policy, Rando will pocket \$750,000 of the up-front payment — an unusually large amount for a deal between a drug company and an academic investigator. Rando's laboratory, his department and the Harvard Medical School each get the same amount.

Isaac Kohlberg, who heads technology development at Harvard, says the agreement shows a new willingness on the part of pharmaceutical companies to reach further 'upstream' into academic labs in search of promising drugs. "Three or four years ago, if a

project was not in phase II [of human clinical trials], large pharma was not interested," says Kohlberg. "Now, when a project has established proof of principle; the mechanism of action is clear; and you have results in animal studies, they will definitely be interested in acquiring rights."

Merck, Harvard and Rando first started talking last spring. Harvard's technology-development office was already in contact with half a dozen companies about Rando's work. Merck, in the meantime, had come across Rando's papers (see "The science behind the deal"). The company, long involved with ophthalmologic drugs, was drawn to Rando's work on visual-cycle inhibition as an approach to preventing AMD.

"Right from the first conversation, there was a meeting of the minds," says Gould. Soon, Merck's top eye-drug developers, based in West Point, Pennsylvania, made the trip to Boston to meet with Rando and Harvard tech-transfer officials. After several months of pounding out details, the deal was finally signed in February.

The Rando agreement is one of only a handful forged between Merck and academics. These are far outnumbered, so far, by its deals with biotech and drug firms. But the balance may be shifting. "This type of partnership really does represent an important part of our strategy moving forward," says Gould. "It's a harbinger of things to come."

1. The Eye Diseases Prevalence Research Group *Arch. Ophthalmol.* **122**, 564–572 (2004).
2. Xue, L., Gollapalli, D. R., Maiti, P., Jahng, W. J. & Rando, R. R. *Cell* **117**, 761–771 (2004).
3. Gollapalli, D. R. & Rando, R. R. *Proc. Natl Acad. Sci. USA* **101**, 10030–10035 (2004).
4. Maiti, P., Gollapalli, D. & Rando, R. R. *Biochemistry* **44**, 14463–14469 (2005).
5. Maiti, P. et al. *Biochemistry* **45**, 852–860 (2006).

THE SCIENCE BEHIND THE DEAL

Robert Rando has spent the last 20 years studying the visual cycle, the complex biochemical pathway by which cells in the retina generate the electrical impulse that gives rise to vision.

Damage those cells — particularly the macula cells in the centre of the retina, which are responsible for high-acuity colour vision — and a characteristic pattern emerges: central blindness. Damage them slowly and chronically, by the build-up of toxic by-products of the visual cycle called lipofuscins, and you

get age-related macular degeneration (AMD). This is a progressive blindness that affects more than one in ten white Americans over the age of 80.

Rando and colleagues first characterized the visual cycle *in vitro* in 1986. After others identified lipofuscins and proposed that they were the toxic culprit leading to cell death in AMD, Rando began a search for visual-cycle inhibitors that he hoped would slow the production of lipofuscins and thus delay or prevent the disease.

To that end, he focused on RPE65, a protein that is essential to the cycle and is found almost exclusively in retinal cells. In two papers in 2004, Rando and colleagues demonstrated that RPE65 is involved in a rate-limiting step in the visual cycle^{2,3}. Combine that with the fact that it occurs almost uniquely in the eye, and you have an apt drug target.

By the time the deal with Merck was being fleshed out, Rando already had in hand data for two more papers in *Biochemistry*. The first, published last October⁴,

defined the specific binding patterns of RPE65. In the second, which appeared in December⁵, Rando and colleagues at Columbia University in New York gave small molecule antagonists of RPE65 to mice that had been genetically engineered to produce excess lipofuscins.

The inhibitors, made in Rando's lab, completely blocked the formation of the most common form of lipofuscin found in the eye. That was enough to persuade Merck to pony up that multi-million dollar upfront fee. **M.W.**

P. PARKER/SPL

Funding: income criterion would hit postdoc careers

SIR — I was surprised to read your Editorial “Brown’s budget briefing” (*Nature* **440**, 581; 2006), backing the UK government’s recent suggestion that external research income should replace much of the current Research Assessment Exercise (RAE). An obvious problem with this proposal is that the RAE affects not only the funding of individual departments, but their hiring and firing policy too. Why would departments ever appoint postdoctoral researchers to junior lectureships if the RAE-replacement was a metric based on external research income? Whatever the quality of their research up to that point, postdocs only become eligible to be principal investigator on research council grants when in a permanent post. An already difficult career step for many postdocs would only get harder.

Phil Bland

Department of Earth Science and Engineering, Royal School of Mines, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

Funding: proposals ignore mentoring and teaching

SIR — Christopher Exley, in Correspondence (“Funding should recognize outcome, not income” *Nature* **440**, 1112; 2006), rightly criticizes the proposal that past research income should be used to assign future funding. But he falls back on the equally simplistic idea that “the only significant research outcome is the science and, primarily, its communication through publication”.

Mentoring and teaching, both of students and the public, have long been the poor relations of publication and funding in the eyes of many researchers. For example, the current strike by UK university lecturers involves the suspension of teaching duties, rather than the suspension of either grant applications or manuscript submissions. Although a boycott of grant applications would, arguably, hit institutes harder than a teaching boycott, lecturers can suspend lecturing at minimal cost to themselves because academic preferment is rarely dependent on the ability to educate.

Given the UK’s shortage of science and engineering graduates, this under-appreciation of education is a terrible and preventable waste. But there is little incentive to mentor in the current UK system and few people, scientists included, do something for nothing. It would be an opportunity missed if more emphasis were not placed on developing the potential of young scientists for the next Research Assessment Exercise.

How many scientific outcomes are more significant than inspiring and training the generation which is to replace us?

John Bothwell

Marine Biological Association of the UK, The Laboratory, Plymouth PL1 2PB, UK

Funding: income is already dependent on outcome

SIR — Christopher Exley, in Correspondence (“Funding should recognize outcome, not income” *Nature* **440**, 1112; 2006), criticizes the use of research income as a metric of science somewhat unjustly. As someone involved in the review of research grants in the United Kingdom, I have found that research income — at least in the case of government research councils, and probably also for charities and industry — is highly dependent on previous and proposed research outcomes. Funding proposals require a brief biography listing previous outcomes as well as a list of expected outcomes and dissemination strategies, both of which reviewers are asked to assess.

Perhaps this is why previous runnings of the Research Assessment Exercise (RAE) have shown a strong correlation between research income and RAE score, at least in disciplines such as science, which require substantial grant income. Although this correlation is not perfect, the time cost to the UK research community of conducting the RAE is not justified by the difference. Further, interdisciplinary researchers who contribute to more than one literature will no longer be disadvantaged.

The main concern is that comparison must be done on a sufficiently narrow scope so that different branches of disciplines do not lose political clout if their research is relatively inexpensive. For example, if comparisons were made at the full departmental level, biology departments would be rewarded for dumping their field researchers in favour of molecular or cell biologists, and computer science departments for dumping theory of computation in favour of robotics.

Joanna Bryson

Department of Computer Science, University of Bath, Bath BA2 7AY, UK

Technology transfer is all about bringing in money

SIR — Your Editorial about technology-transfer offices, “More than the money” (*Nature* **440**, 845–846; 2006), is, in my opinion, dangerous and wrong.

It is dangerous because, at a time when governments are investing in science with the

reasonable hope that new knowledge can lead to economic advantage, it is ammunition for critics to suggest otherwise.

It is wrong both in principle and in practice. The chemistry department at Oxford has contributed some £80 million (US\$150 million) to the central university over the past ten years. Actual realized gains from the departmental spin-offs come to more than £40 million, with about £20 million of unrealized gains in quoted companies and a further batch of holdings in private companies. Of these, one has raised several millions in a pre-flotation round, another has a multimillion-pound contract with a major pharmaceutical company, and a third has a huge European Union grant. This has been achieved without any pressure on academics to modify their research or even to push people into commercialization.

Many technology-transfer offices are frankly not very good. They should not be allowed to pretend that bringing money in to their universities is not one of their major roles.

W. Graham Richards

Central Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QH, UK

Debye's wooden response undercut Nazi orders

SIR — I read with chagrin your News in Brief story “Dutch universities ditch reported Nazi collaborator” (*Nature* **440**, 139; 2006), stating that two Dutch universities have rescinded the recognition previously accorded to Peter Debye. I was a junior member of his chemistry department at Cornell University and I remember two stories in particular that make it hard to believe Debye had Nazi links.

Debye told us about one of the buildings at the Kaiser Wilhelm Institute for Physics, where he was director, that had the name of the previous director, Max Planck, chiselled in stone above the entrance. When ordered to erase the dedication to Planck (who had tried to help Jewish colleagues), Debye had workmen cover it with a large wooden board. That way, when visitors asked about the unseemly aberration, he would feel obliged to answer.

He also described how he was finally convinced to flee Germany when a direct order came from Hitler to put all the institute’s resources into the development of the atomic bomb. Carrying only their suitcases, Debye and his wife travelled to the Netherlands, ostensibly for a lecture tour, but they never returned. Instead they went on to London and eventually to America.

Harvey Posvic

Department of Chemistry, Loyola University Chicago, 1068 West Sheridan Road, Chicago, Illinois 60626, USA

COMMENTARY



S. PEMBERTON

Europe pays the price for spending less

Europe's contribution to the global advancement of science and the promotion of learning is in decline. Better funding of universities and research institutions is needed to reverse this trend, argues **Christopher Patten**.

The British Surgeon General in the Crimean War, responding to complaints about the quality of the medical services for which he was responsible, argued that they would have been perfectly adequate had it not been for the war's casualties. I am reminded of that insouciance when Europe's political leaders bemoan the European Union's growing problem of international competitiveness. The difficulty they have in relating cause to effect is, I suppose, an example of their difficulty in grasping the scientific method.

Cost of knowledge

I had a fairly orthodox career in national politics, including one year as the education minister. I spent five years in Asia, presiding over a colony in which the provision of higher education grew exponentially — underpinning the transition from a low- to a high-value-added economy and providing a rite of social passage for the offspring of the mainly immigrant community. I spent a further five years as a European commissioner, considering, sometimes morosely, the gap between Europe's global aspirations and its ability to live up to them. I am now a chancellor of two fine universities, one of which is among the greatest in the world, with an unrivalled brand and image. Finally, I chaired a steering committee for the European Research Council, suggesting ways

in which that body should operate, free of bureaucratic or political interference.

All that gives me what others might call prejudices, but which I would naturally argue are informed opinions. I am not myself an academic, nor an administrator of educational or research programmes. Nevertheless, I believe my views reflect a growing opinion in Europe, that much of its higher-education system is in severe difficulties; the research base is threatened; many of the best researchers are being lost; and there is more competition in the knowledge business. There are serious consequences for Europe's future as an economy and a civilization.

All this may surprise those who believe what is said by heads of state and finance ministers at European councils. Since 2000, the European Union (EU) has been committed to a strategy to become the most competitive knowledge-based economy in the world by 2010. To accomplish this heroic objective, reforms to higher education and research and development (R&D) have been proposed, and a number have even been implemented, with some countries doing much better than others (http://ec.europa.eu/invest-in-research/national/national_reform_en.htm). But overall, the picture is not looking good, and some of the reasons are perhaps understandable. Opening up the market in public utilities,

reforming labour markets, facing down protectionist pressures at a time of low-ish economic growth — all could be said to pose political difficulties. But what is the problem with investing more in knowledge? The biggest problem by far is simply a lack of money from governments (or other sources), together with inadequate corporate investment in R&D.

Lower education

The consequences of the present situation can be described in both conventionally utilitarian and liberal terms. The utilitarian argument is that we live in a knowledge-based economy — a cliché that is true, up to a point. I do not seek to argue that there is a direct correlation between spending on higher education and economic growth, nor between investment in R&D and successful innovation. Nevertheless, it is surely unarguable that science is a principal determinant of the wealth, creativity and well-being of society. Yesterday's good science — answering the 'whys' — added to today's technology — working out the 'hows' — has provided much of the quality and standard of our lives today. We are now living off what has been achieved in the past by researchers, technologists and entrepreneurs. What will Europe hand to the next generation?

Turning from concept to statistics, spending

on tertiary education per student (including R&D) in US\$ is just over 9,000 for France, just under 11,000 for Germany and just under 12,000 for the United Kingdom (OECD 2005 indicators). Austria, Belgium, Denmark, the Netherlands and Sweden do better, yet by comparison, the figure for the United States is heading towards 26,000. The percentage of gross domestic product (GDP) spent on tertiary education (2002 figures) is 1.1% in France, Germany and the United Kingdom, and 2.6% for the United States (1.2% from public funds and 1.4% from private). In other words, the public sector in the United States is more generous than the public and private sectors combined in Europe's three largest economies. Universities are the principal incubators of research. If universities are undervalued, what are the prospects for research?

What else do the figures tell us? The international league table of university performance of London's *Times Higher Education Supplement* does not reflect perfect methodology (how could it?), but uses various weighted factors covering peer review, recruiter review, percentage of international faculty, percentage of international students, faculty-to-student ratio and research impact measured by citations per faculty member. It concentrates on large general universities and excludes institutions that do not teach undergraduates and those with fewer than 5,000 papers in the citation index.

In the *Times* 2005 rankings, there are three European institutions — Oxford, Cambridge and École Polytechnique — in the top ten; the rest are all from the United States (see Table). In the next ten are the London School of Economics and Imperial College, London, the rest being US institutions plus Beijing, Tokyo and Melbourne. The next European university is the École Normale Supérieure at 24, and the first German university is Heidelberg at 45. Altogether in the top 50 there are 13 European institutions. A similar list drawn up by Shanghai Jiao Tong University, China, in 2005 is more depressing for Europe: there are two European universities in the top ten, and nine in the top 50, of which one is Swiss and none is German.

What does this tell us? Not long ago, Europe came way ahead in such lists of academic attainment. After all, it pretty well invented universities, such that the traditional liberal arts universities in the United States drew their inspiration from the British university system, and the great US research universities were founded on the Humboldt (German) model. Is it not a paradox that during a period in Europe of unparalleled prosperity and stability, the priority given to universities in, for example, the allocation of public resources, has sagged so pitifully?

Public controversy focuses on health and on welfare entitlement spending. Are universities victims of competitive populist politics? Europe spends very little on its own security in comparison to the United States. Is it not far

more culpable that it spends so much less on knowledge and learning? After all, there are political objections, right or wrong, to more spending on defence, but what is the argument against spending more on universities? How can Europe be so condescending about US culture when that country spends twice as much on knowledge, its transmission to students, and its acquisition, as Europe? Europe is allowing its culture to wither on the vine.

Brain drain

Two other statistics have made an impression on me. In the first three decades of the twentieth century, nationals of France, Germany and Britain each won more Nobel prizes in science and economics than the United States, which collected just 3% of the total. Since 1970, the United States has picked up almost 60%, more than the whole of Europe combined. The

TOP 20 UNIVERSITIES IN 2005

Rank	Name	Country
1	Harvard University	US
2	Massachusetts Institute of Technology	US
3	Cambridge University	UK
4	Oxford University	UK
5	Stanford University	US
6	University of California, Berkeley	US
7	Yale University	US
8	California Institute of Technology	US
9	Princeton University	US
10	École Polytechnique	France
11=	Duke University	US
11=	London School of Economics	UK
13	Imperial College London	UK
14	Cornell University	US
15	Beijing University	China
16	Tokyo University	Japan
17=	University of California, San Francisco	US
17=	University of Chicago	US
19	Melbourne University	Aus.
20	Columbia University	US

European tally has looked more respectable in the past five years if measured by the winners' countries of birth, as the United States has greatly benefited from a brain drain — and still does.

What will the figures look like in future? Winners of Nobel prizes are usually rewarded for work done after the completion of their doctorates. A decade ago, half the young Europeans who had completed doctorates in the United States came home afterwards; yet now, only a quarter do so. None of us should want to put up barriers against the movement of academics; scholarship should know no boundaries. European academic institutions need to be able to attract the best to return.

The other statistic commonly used to measure academic achievement is citations used in academic papers. Today, the United States accounts for half of these. Yet many of these papers are written by foreign researchers — mainly Asians and Europeans — based in US academic institutions.

Although I am wholly in favour of expanding educational opportunity, it should not be at the expense of lessening the quality of what is being offered. An increase in quantity is to

increase social inclusiveness and to broaden the skills of the workforce. In Britain, at least, university study is now seen by those undertaking it as a necessary step to getting better-paid employment. A university degree is the best indicator of future income (unless you decide to become an academic!).

Because governments have been unwilling to raise spending in line with the expansion of the number of students, they have concentrated on the sort of objectives that would make more sense in running profit-making enterprises. Per capita costs become the issue, not the quality of the learning experience, and those who are taught protest less than they should because the system has become more a matter of credentialing — as the economist and moralist, Jane Jacobs, has argued (J. Jacobs, *Dark Age Ahead*, Random House, New York; 2004) — rather than learning. This contributes to high drop-out rates in some countries, and absurdly long periods taken to graduate in others.

Governments marvel at how, in comparison with, say, the expansion of healthcare, they have managed to open up opportunities in higher education exponentially without a concomitant expansion of spending. They find the answer in allegedly increased productivity. But what that means is a downward pressure on academic salaries, degraded facilities and a devalued learning experience. The well-intentioned expansion of education is threatening its quality.

Research standstill

How much does all this matter to performance in research, development and innovation? First, it would be more than surprising if the condition of the principal agents for promoting both advanced learning and pioneering research did not have an impact on the broader research scene. Second, underinvestment in research in universities is matched by underinvestment in business. Corporate Europe does much less than corporate America and Japan. As a percentage of GDP, R&D in Europe has stagnated for ten years at round about 2%, while it has been growing in the United States. There, spending as a proportion of GDP is almost half as much again as the European figure. There is a bigger gap when Europe is compared with Japan. Much of the gaps are accounted for by higher funding by the business sector: in the United States about a third of government R&D funding goes to support business R&D, twice the European figure.

Europe's corporate sector's performance is increasingly lacklustre: last year's UK government figures showed a 7% increase in R&D spending in the United States and Asia, and 2% in Europe. Greater R&D spending does not necessarily mean more innovation, as other factors are clearly involved, not least the quality of the investment in research and perhaps, above all, market competition. But surely in



The European Commission is attempting to increase investment in research and development, with mixed results.

some of the frontier industries — pharmaceuticals, biotechnology, IT hardware and software, for example — R&D spending is vital to innovation. WalMart or Tesco may not have to spend much on research in order to innovate. Glaxo and Nokia are in a different category. Moreover, it may be that the simplest way of showing the corporate benefits of R&D is the share portfolio of R&D-intensive companies.

The competition is hotting up elsewhere. The number of scientific papers being produced in Asia is rising fast, for example in nanotechnology journals. India's institutes of technology grow in quality, while the country already produces 260,000 engineers a year. China's spending on R&D has trebled in seven years and is predicted to hit 2% of GDP by 2010.

So what can Europe do? It is important both to recognize the scale of the problem, and to understand that it should be easier to solve than some others. Europe starts with a huge advantage in terms of standard of living and quality of life in comparison with much of the rest of the world. Europe is not only prosperous, but has great universities and research institutes, with long histories, civilized traditions and dedicated staff. All that they lack, on the whole, is more generous support. Similarly, Europe has many world-class companies.

Some things require action at the national level; some at the European level. EU member states should be encouraged to give a greater priority to higher education and R&D, with a scoreboard of their performance across the EU. To the extent that governments do not require students to pay for tuition, the gap can only be made up by the taxpayer, the corporation or the philanthropic giver. If governments spent more, that might well encourage greater corporate investment both in universities and in their own laboratories. More philanthropic support of universities should be encouraged: only two universities in Europe would get into

a list of the top 150 in the United States for the size of their private endowments.

The primary responsibility for improvement lies with national governments, but there is also plenty that can be done at the European level. The coverage of the Bologna process — a scheme launched in 1999 to align and harmonize degree standards and the quality of courses and university departments in Europe — is patchy and needs to be broadened. The lack of a European patent continues to load additional costs on to European inventors in comparison with US competitors. This is a subject that has been debated for three decades. Each time the negotiations founder on the language issue: if Europe is serious about competitiveness, this can easily be resolved.

Action underway

At least the European Commission seems to understand the nature and worrying scale of the problem. It has tried to do three things — one unsuccessfully, one with every possibility of success, and one that I cannot quite fathom.

First, it has tried (as recommended in 2003 by the Sapir report) to reorient the European budget towards what should be its priorities — competitiveness, growth and R&D — and away from agricultural-led priorities set in the 1950s. It has not had much success.

On the second issue, the Commission has been much more successful. A group, chaired by myself, proposed among other things the establishment of the European Research Centre — an equivalent to the US National Science Foundation — which is turning from a caterpillar into a butterfly. It will seek to fund frontier research in top-class departments and institutions, and will be run by scientists, academics and researchers, not bureaucrats or politicians. It will be driven by peer review so that grants will go to the best, not distributed on the basis of *juste retour*. Political interfer-

ence should be minimal — here Europe has the priceless advantage over the United States that scientific research has not been assaulted by the anti-enlightenment forces of unreason except in the case of animal research.

High-quality, curiosity-driven research can be even more commercially significant than proprietary science — think of lasers and DNA. Knowledge is not simply another commodity in the market place. Yet while encouraging greater cooperation between industry and universities, research must not be skewed almost exclusively to subjects that interest industry rather than the public. The European Research Council is on its way, with the only major matter still to be resolved being the size of its budget.

That is where the third Commission proposal jars. It has suggested the establishment of a European Institute of Technology (EIT), presumably to try to replicate the Massachusetts Institute of Technology and its impact on its regional economy in the north-east United States. But is the EIT a real institution or is it a virtual body, promoting yet more networking? Is it meant to create more technology and innovation clusters such as those that already exist in European countries? I believe we do not require one new institution, but much better funding of some of the existing universities and research institutions. Finally, what guarantee is there that the money for the EIT will not come from what would otherwise be spent through the European Research Centre on research of peer-reviewed excellence? I cannot see how the EIT — however skeletal its structure — will do other than take money from the funds that would otherwise go to existing institutions. ■ Lord Patten is Chancellor of the Universities of Oxford and of Newcastle, UK. This text is an edited version of a speech by the author to the Academy of Technologies in Paris, France, on 28 February 2006.

BOOKS & ARTS

From cell to organism

How genetics unlocked the mysteries of development.

Coming to Life: How Genes Drive Development

by Christiane Nüsslein-Volhard
Kales Press: 2006. 176 pp. \$29.95

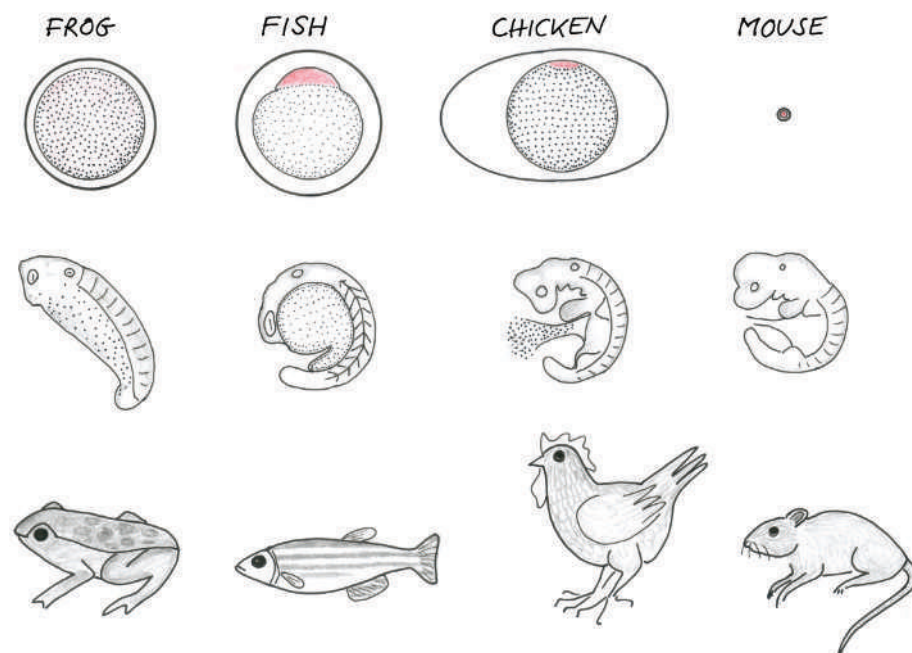
Alfonso Martinez Arias

It is easy to forget how little was known about the development of living organisms just 30 years ago. Today we have whole genome sequences of many species that can be compared, endless patterns of gene expression that are just beginning to be pieced together, and technologies such as RNA interference that allow us to do genetics at amazing speeds and depths. Together these provide an opportunity to understand the mechanics of life and the details of how plants and animals are formed. And yet, at the beginning of the 1970s, the development of an organism from a fertilized egg was still a marvel with a hidden secret.

Yes, there were hints that the secret would soon be unlocked. The uncovering of genetic regulatory circuits in bacteria had famously led Jacques Monod to state that what was true for *Escherichia coli* would also be true for the elephant. However, despite these high hopes, the secret remained out of reach. In the 1980s there was a watershed. It came not from the application of biochemistry to the problem, which had been tried before without much success, but from applying a judicious mixture of genetics and molecular biology to two invertebrates: the nematode *Caenorhabditis elegans* and the fruitfly *Drosophila melanogaster*.

Christiane Nüsslein-Volhard has played a central role in the process of discovery that has enlightened our understanding of how one cell becomes an organism. Her contributions have the virtue of having explored the issue both in invertebrates, with her pioneering studies of the early development of *Drosophila*, and more recently in vertebrates, through her work on the developmental genetics of the zebrafish *Danio rerio*. It is a range allowing her a rare insight into the development of organisms, and this permeates *Coming to Life*, an account of where we stand in our understanding of the process of embryonic development.

The content of the book ranges from history to evolutionary biology, and includes a crash course in molecular biology and insights into the molecular embryology of different organisms. Surprisingly, all this ground is covered



Hand-drawn illustrations bring a personal touch to Christiane Nüsslein-Volhard's book on development.

in 140-odd pages of main text with a style that is direct, simple and easy to read. The book is carefully produced and the illustrations are hand drawn and full of personal insight.

The thread of the book is the importance of genetic analysis as an essential tool to understanding the development of organisms, and the principles that have been derived from this. A second important message is the sense of unity of mechanism that underlies the diversity of animal form. Naturally, *Drosophila* is central to the book, and we have clear accounts of the role that genetics has played, and continues to play, in uncovering developmental principles.

Nüsslein-Volhard highlights some of the barriers that have been broken in the process of establishing the existence of these principles. Nowhere is this better illustrated than in the uncovering of informational gradients and their role in pattern formation. Classical embryological experiments hinted at their existence, but brute-force biochemistry had failed to identify such factors. It was a genetic approach that led to the breakthrough with the *bicoid* gene and its role in early *Drosophila* development. Genetics also provided the logic behind patterning along the axes and

revealed that different organisms represent different outputs of a conserved molecular kit made up of transcription factors, such as the Hox proteins, and signalling molecules that allow cells to communicate with each other. Nüsslein-Volhard's book establishes the link between *Drosophila* and vertebrates, and discusses the basic facts of human embryology as another, albeit special, case of vertebrate embryology.

For the most part, the book is an account of a mature field of study with some personal touches. The mostly objective mould is broken in the final chapter, in which Nüsslein-Volhard tackles several sensitive issues concerning the nature of life and its implications for human welfare and reproduction, and the potential of stem cells and genetic therapies. She is not afraid of giving her opinion: "Although the criteria used for the definitions rely on biological events, they are not a scientific but a moral issue. Dignity, right of life and protection are not biological but moral categories ... these issues should be decided not by scientists but by our society as a whole through our political representatives." The function of the scientist, as most of us will agree, is to find out how things work and to provide information

that can be used to help others take decisions. Our role is not to cast shadows or to mediate those decisions. This chapter also makes clear what is possible and what, at the moment, is a utopia.

Coming to Life is said to be aimed at “those who are curious and who would like to understand the process of life a little better without having to deal with highly specialized knowledge”. The final chapter in particular makes

one wonder whether reading this book would not also benefit politicians, social activists and journalists who ponder issues such as cloning and stem-cell or genetic therapies in the public realm. They might find the science rather challenging in places, but it would certainly give them a perspective for their judgements. ■

Alfonso Martinez Arias is in the Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK.

geographers had reverted to the earlier idea, proposed by the Greek scholar Eratosthenes, that the world consisted of inhabited land surrounded by the ocean, as opposed to Ptolemy's view in which oceans appeared as giant lakes. For example, the map shows the Arabic error of considering the Malaysian peninsula to be a vast elongation of southern China, but separated from Africa, as seen in al-Umari's map. An even older version of the map discovered by Joseph Needham corroborates Sezgin's reconstruction.

Once the Ma'munic improvement on Ptolemy was correctly reconstructed, it led to a search for other geodetic improvements, both in methods and in results, by Islamic mathematical geographers and cartographers. Sezgin documents in great detail how the Ptolemaic ‘long Mediterranean’ was shortened almost to its length today by fixing the positions of Baghdad and Toledo; how the Caspian Sea gradually acquired its correct outlines; how, by a combination of dead reckoning and trigonometry, the Arabs managed to get to grips with longitude at sea; and finally, how the distance between Ghazna and Baghdad was measured with astonishing precision in the eleventh century using astronomical observations and spherical trigonometry.

Sezgin then documents the influence of these achievements in Europe. A number of Western authors, including Marco Polo, Dante, the astronomer Johannes Kepler and the cartographer Nicolas Sanson, seem to have been informed by Arabic-Islamic geography. Sezgin shows how the Muslim maps led to the portolans, as they had to Portuguese maps of the Indian Ocean. He documents and discusses in great detail the information that Muslim geographers gathered on Asia, the

Charting an Arabic course

Mathematical Geography and Cartography in Islam and their Continuation in the Occident, Vol. I. Historical Presentation

by Fuat Sezgin, transl. Guy Moore & Geoff Sammon

Institut für Geschichte der Arabisch-Islamischen Wissenschaften, Johann Wolfgang Goethe-Universität: 2005. 636 pp. € 175

A. M. Celâl Şengör

Few books in the historiography of science have single-handedly created a revolution in their subject, but this one just might. It is the result of a 15-year effort by the renowned Arabist and science historian Fuat Sezgin. Originally, the book was to be just another volume in the author's *History of Arabic Literature* series, begun in 1967, in which a scholarly historical introduction is followed by a list of authors and critical assessments of their works. *Mathematical Geography and Cartography in Islam and their Continuation in the Occident* retains this format, but the usual introduction is expanded into two massive volumes of ‘historical presentation’, plus a lavishly illustrated atlas with 209 maps (many in colour). It was originally published in German in 2000, but an English translation of Volume 1 is now available and Volume 2 will follow shortly, along with the final ‘authors’ section of the German original.

While working on the book, Sezgin found that conventional treatments of the history of mathematical geography and cartography had failed to take into account nearly half of the available material, and it is the narration and analysis of this that fills the bulk of these volumes. He first became aware that something was seriously amiss when he heard about the sudden appearance of some portolan maps, or navigational sea charts, from around 1300, and some remarkably accurate longitude and latitude data in European maps of parts of Asia, dating from times when no European had been there to take any measurements.

A clue to the puzzle lay in the discovery of an annotated version of *Masalik al-abşar fi mamalik al-amşar* by Ibn Fadlallah al-Umari, a representation from 1340 of a lost map of the world prepared by the geographers of

the Abbasid caliph al-Ma'mun in the eighth century. The version by al-Umari was drawn on a pre-existing graticule with a globular projection, and not vice versa as had previously been assumed — I can vouch for this, having checked the original myself. This helped Sezgin interpret the data on geographical coordinates that originally accompanied the Ma'munic map. Sezgin found the data preserved in the book *Sûrat al-ard* from the early ninth century, which has commonly, but mistakenly, been considered to be an original work of Abu Ja'far al-Khwarizmi. It is now thought to be a simple register of Ma'munic location data.

Sezgin argues that the collection of all these data necessitated new fieldwork and must have required a large group of observers, rather than being revisions by one hypothetical Syriac author of a Ptolemy translation, as Hans von Mzik had proposed. Sezgin's reconstruction of the Ma'munic map, put together using the only two preserved versions of its coordinate data, shows that the Ma'munic



An Arabic world view: Ibn Fadlallah al-Umari's map from 1340 contains data from the eighth century.

TOPKAPI PALACE MUSEUM

Indian Ocean, the Malay Archipelago, the Mediterranean and Africa. He also details the new geodetic methods they discovered and the numerous coordinate tables they generated, which influenced European maps well into the eighteenth century.

The abundance of material brought together in this book is truly awesome. The three handsomely produced volumes are thoroughly indexed. With so many novelties, Sezgin's book will no doubt be controversial. But therein lies an added value: it will provoke an even more

thorough search for additional data on the history of a subject that had long stagnated. ■
A. M. Celâl Şengör is in the Department of Geology, Istanbul Technical University, and the Eurasian Institute of Earth Sciences, Ayazaga, 34469 Istanbul, Turkey.

THEATRE

Would the real Mr Feynman...?

Clever Dick

written and directed by Crispin Whittell

At Hampstead Theatre, London, until 17 June
www.hampsteadtheatre.com

Richard Webb

"Confusion. See, confusion ain't a sickness I suffer from. I take my strength from the simple love of my country and faith in my God." Self-doubt does not cloud the horizons of 'Fat Man' Sergeant Whitey Stokes, Counter-Intelligence Corps, on the evening of 17 June 1945. The Red spy he thinks he's interrogating — the physicist Richard Feynman — on the other hand, is very confused. He hasn't slept for 58 hours. On his way to Los Alamos, he took the wrong turning at Albuquerque because of a congenital inability to tell left from right. Mistaking him for her beau 'Little Boy', the 18-year-old receptionist of his hotel has just tried to go to bed with him.

Thus, in *Clever Dick*, playwright-director Crispin Whittell lays out the ingredients for a classic farce. But the sobriquets of two characters reveal a deadly serious setting: Little Boy and Fat Man were the nemeses of Hiroshima and Nagasaki. One month after Whittell's imagined encounter in a New Mexico hotel room, the first atomic bomb is to be tested.

Hence Feynman, played by Adrian Rawlins with a nervous, nerdy energy that contrives to make his slight build seem lanky. His busy hands conduct to a score with some diverting passages ("So he knows a lot of Communists. I know a lot of women, but that doesn't make me a woman") and frenetic, slapstick moments — topped by a spectacular re-entry for Fat Man, amply portrayed by Corey Johnson, that literally brings the house down.

But ultimately this farrago is, like that ceiling, a false plasterboard mishmash. Its profound periods are not deep enough nor its farcical moments sustained enough to carry the audience, which is only truly captivated by a moment of pathos towards the end when Feynman reveals the real reason for his confusion: the death, the evening before, of his wife from tuberculosis. Science enters in the second half only in a shampoo-ad-style 'here's the science bit' as a fig-leaf for the play's intellectual flaccidity. Feynman, rigorous scientist and, above all, great communicator of modern physics, would, one suspects, not have been pleased.

Indeed, what is Feynman doing here? He is pitched — as a participant in the Manhattan project that developed the atom bomb — as the voice of nuanced moral doubt against the dumb certitude of the gung-ho Stokes with his unshakeable faith in God and the gun (an intended contemporary resonance). But this is a role that Feynman never played: his principal extramural contribution to the Manhattan project lay not in breast-beating but bongo-beating and the apocryphal cracking of every safe containing nuclear secrets on the Los Alamos site. He was ever, in Freeman Dyson's words, "half genius and half buffoon". As he said in a 1981 interview: "What I did — immorally, I would say — was to not remember the reason that I said I was doing it, so that when ... Germany was defeated, not the slightest thought came to my mind ... I simply didn't think, okay?"

Michael Frayn's *Copenhagen* left a great appetite for plays that deal with the interface of science and morality. This is Whittell's second foray onto such territory, following his generally well-received 2003 play *Darwin in Malibu*. The safe-cracking, wisecracking womanizer Feynman is a headline-grabbing front man for a second attempt. In this case the result is, charitably put, confused. ■

Richard Webb is the physical-sciences News and Views editor at *Nature*.



Fat Man (Corey Johnson) and Feynman (Adrian Rawlins) in *Clever Dick* at the Hampstead Theatre.

K. PATTISON

The eye of the beholder

On Seeing: Things Seen, Unseen and Obscene

by F. Gonzalez-Crussi

Overlook/Duckworth: 2006. 224 pp.
\$24.95/£14.99

Richard Gregory

Frank Gonzalez-Crussi is a pathologist, well versed in how we see, although here he focuses on what we see, and on what, for obscure reasons, are rejected by eye and mind. This disportment in art and medicine is rewarding thanks to the power of the writing to carry the reader safely to "things seen, unseen and obscene". These include the body, viewed from inside and outside with a medical gaze; French literature and history (with a hilarious account of the start of the Revolution triggered by

a pair of royalist voyeurs); belief; and sin. A principal theme is the attraction and repulsion of the anatomical focus of femininity to the male eye, which induces a fascinated fear, rather like standing on the edge of cliff.

Gonzalez-Crussi suggests reasons for why anatomy was so slow to develop as a science, including the inability to make effective use of detailed knowledge before anaesthetics. But has a lack of rapid use impeded other sciences? If it is inhibition associated with prurience that held back the study of anatomy, it is unfortunate that this book was not available centuries ago. It can be enjoyed for its broad learning, and its broad humour. ■

Richard Gregory is in the Department of Experimental Psychology, University of Bristol, Bristol BS8 1HH, UK.

CHEMISTRY

Synthesis with a twist

Harry H. Wasserman

Amide bonds underpin protein stability, but distortion from their preferred geometry makes them reactive. An archetypal twisted amide has at last been made, and its chemistry will be revelatory.

On first appearances, 2-quinuclidone is a small, uncomplicated molecule compared with the many complex systems made by organic chemists in recent years. Yet the synthesis of this compound by Tani and Stoltz described on page 731 of this issue¹, and the study of its chemical behaviour that this will enable, represents a milestone in our understanding of the properties of a special type of bond — the twisted amide.

The amide is one of the most fundamental functional groups in chemistry. It is the link between amino acids in peptides and proteins, and thus is key to the structure of many biological systems. Amides are surprisingly robust compared with structurally related derivatives, and it is believed that this linkage gains stability from electron 'delocalization'. The amide group can be thought of in terms of two structures (Fig. 1a), one of which is apolar, and the other dipolar. These two forms — known as resonance structures — differ in the location of their double bond, and represent two theoretical possibilities of where the associated electrons could reside. In reality, the electrons in the double bond are spread (delocalized) across the amide group, and behave as a hybrid of the two theoretical distributions.

For the dipolar form to contribute to the linkage, the chemical bonds around the amide have to lie in the same plane, in order to satisfy the geometrical requirements of the carbon–nitrogen double bond². In circumstances where such coplanarity cannot be achieved without distortion of the structure, as would be the case with 2-quinuclidone (Fig. 1b), stabilization of the amide bond by electron delocalization is inhibited. Such twisted amides demonstrate unusually reactive chemical behaviour compared with typical amide bonds.

The 2-quinuclidone structure and reactivity problem has a fascinating historical connection. In 1941, during his studies on the synthesis of quinine, the young Harvard investigator (and eventual Nobel prizewinner) Robert B. Woodward thought deeply about the reactivity expected of this type of distorted amide³. As one of his first graduate students, I was given

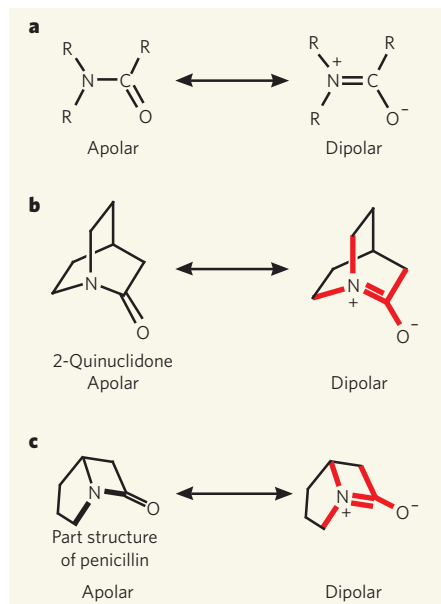


Figure 1 | Electron delocalization in amides.

a, Standard amide bonds (where R is a hydrogen atom, or any general hydrocarbon chain) are a hybrid of two forms, known as resonance structures. The double-headed arrow indicates that the true electronic structure of the bonding lies somewhere between the extremes of an apolar and a dipolar form. The electrons are actually spread out between the nitrogen, carbon and oxygen atoms. This electron delocalization, which strengthens the amide bond, is only maximized if the atoms around the carbon–nitrogen double bond in the dipolar form are coplanar. **b**, For 2-quinuclidone, the geometry of the molecule prevents coplanarity of the relevant bonds (marked in red) in the dipolar form. Electron delocalization is inhibited, so the carbon–nitrogen bond is weak compared with untwisted amides. **c**, In the dipolar form of penicillin, the bonds in red cannot all be coplanar, again inhibiting electron delocalization through resonance, and destabilizing the amide bond.

the problem of synthesizing 2-quinuclidone. Originally, this research was largely of academic interest, but a few years later the expected chemical behaviour of twisted amides became an important issue in the determination of the structure of penicillin.

The massive international effort to produce penicillin during the Second World War required an unambiguous structural determination of the antibiotic, so that a chemical synthesis could be attempted. There was, however, a serious difference of opinion regarding the structure among the investigators at a time

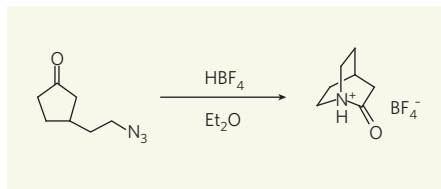


Figure 2 | The preparation of 2-quinuclidone. The precursor molecule undergoes a rearrangement in the presence of tetrafluoroboric acid (HBF₄). Hydrolysis of the product is prevented by the use of a non-aqueous solvent (diethyl ether (Et₂O)). 2-Quinuclidone is basic, so the product forms a salt with the tetrafluoroboric acid.

when modern analytical methods were not available. The question revolved around the marked reactivity of penicillin towards decomposition with water, a process known as hydrolysis. Woodward championed a so-called β-lactam structure, which contained a cyclic amide fused to a five-membered ring² (Fig. 1c), whereas another group, led by Robert Robinson, noted that model β-lactams did not show enhanced penicillin-like reactivity and so strongly favoured an alternative molecular arrangement. Woodward pointed out that the geometry of his suggested β-lactam forced the bonds of the dipolar contributor out of coplanarity (Fig. 1c), inhibiting electron delocalization and leading to ready hydrolysis. In other words, his suggested structure contained a reactive twisted amide. The β-lactam structure was soon accepted, in large measure owing to the logic of Woodward's arguments².

As twisted amides can be an essential design feature of biologically useful molecules, the study of 2-quinuclidone, a model twisted amide which does not exist naturally, is of great interest. There have been a number of attempts to prepare this compound, but no

successful synthesis has been previously described. Tani and Stoltz¹ designed their approach so as to avoid aqueous conditions that would promote decomposition of the product. Accordingly, they focused on a route involving a molecular rearrangement, known as the Schmidt–Aubé reaction (Fig. 2), which is carried out in the absence of hydrolytic reagents. Because 2-quinclidone is a base, and the reaction requires an acid, the product is obtained as a salt.

The authors outline in detail the studies leading to the optimal reaction conditions, and describe the spectroscopic and chemical properties of the product. An X-ray analysis confirmed the highly twisted nature of the

amide. Tani and Stoltz's imaginative, unambiguous synthesis of 2-quinclidone provides important confirmation of earlier predictions regarding its reactivity. This work will provide an invaluable insight into the nature of amides, and concludes a synthetic journey that began nearly 70 years ago. ■

Harry H. Wasserman is in the Department of Chemistry, Yale University, 225 Prospect Street, New Haven, Connecticut 06511, USA.
e-mail: harry.wasserman@yale.edu

1. Tani, K. & Stoltz, B. M. *Nature* **441**, 731–734 (2006).
2. Johnson, J. R., Woodward, R. B. & Robinson, R. in *The Chemistry of Penicillin* (eds Clark, H. T., Johnson, J. R. & Robinson, R.) 440–454 (Princeton Univ. Press, 1949).
3. Wasserman, H. H. *Heterocycles* **7**, 1–15 (1977).

PALAEONTOLOGY

Respect for stromatolites

Stanley M. Awramik

Is it time to stop worrying over whether the ancient structures called stromatolites are of microbial origin? 'Yes' is the answer to emerge from field and lab work on a 3,430-million-year-old marine ecosystem.

Writing in these pages earlier this year¹, Don E. Canfield wrote of an early Earth teeming with a variety of microorganisms. Stromatolites were mentioned as among the possible evidence of microbial life existing 3.5 billion years ago, but only as “probably” of biological origin. What is it about stromatolites that engenders caution when interpreting them as fossils? In their paper on page 714 of this issue², Allwood and colleagues provide good reasons to suppose that such reservations really aren't necessary.

Part of the caution in interpreting stromatolites stems from their nature. They are laminated sedimentary structures, with shapes that range from simple domes to elaborately branched columns, and they range from the millimetre scale to more than 10 metres in size. Modern analogues are known where photosynthetic microorganisms — principally cyanobacteria — are responsible for forming them. Only rarely are microfossils found in ancient examples, but many researchers consider stromatolites to be the products of microbe–sediment interaction, and so to be fossils³.

Confidence in the interpretation of stromatolites as biogenic structures was dealt a serious blow with the proposal that stromatolite structure could theoretically result from abiotic processes in sediment accumulation⁴. To make matters more bewildering, the very definition of stromatolite is contentious^{3,5}. So there has been a long-standing debate about whether these laminated sedimentary structures are indeed indicators of ancient microbial life⁵ — a debate that has intensified with the increased attention being paid to searches

for evidence of life on other planets, and so to the earliest record of life on Earth⁶.

This is the background for Allwood and colleagues' report². They investigated the relationship of stromatolite shape to depositional environment across more than 10 km of nearly continuous outcrop of the 3,430-million-year-old Strelley Pool Chert in Western Australia (Fig. 1). Rather than wade through long philosophical and semantic arguments (theirs are brief and to the point), Allwood *et al.* wade through the evidence. They conclude that the stromatolites are biogenic in origin, and were formed on a 'carbonate platform' — a broad,

relatively flat expanse of carbonate sediments — that was subject to the influence of tides but generally remained submerged. Sea-level rise (transgression) was a major feature of this environment, but small fluctuations — transgressions and regressions (sea-level fall) — also occurred. At times, regression and restricted circulation with the open sea took place, leading to evaporation of water and the formation of evaporite minerals.

Allwood *et al.*² recognize seven distinct stromatolite facies (sedimentary rocks with distinguishing characteristics that correspond to the conditions of deposition). They identify distinctive stromatolite shapes in each facies, and the supplementary material for their paper presents a dazzling array of images and information. This is not the first report on stromatolites from this formation^{7,8} and it probably won't be the last.

What is so striking about their scenario is that it is so 'normal' — in the sense that transgressive, tidally influenced carbonate platforms with stromatolites are common in the geological record^{9,10}. Stromatolites are also known to form in similar environments today. Given this context, and combined with other features, Allwood *et al.*² argue that the most likely interpretation is that the stromatolites are biogenic. Among those features are the seven distinct stromatolite types and their facies (see Fig. 1 of the paper² on page 716); the geometries of the stromatolites; and the fine details of the sedimentary textures of the stromatolites themselves and the regions between them (the 'interspace' areas).

The sediment-grain compositions and textures of the stromatolites cannot be explained exclusively by mechanical processes. For example, the laminae of the 'complex cone' structures are not of the equal thickness that would indicate abiotic precipitation. And the sedimentology of the cones differs from that

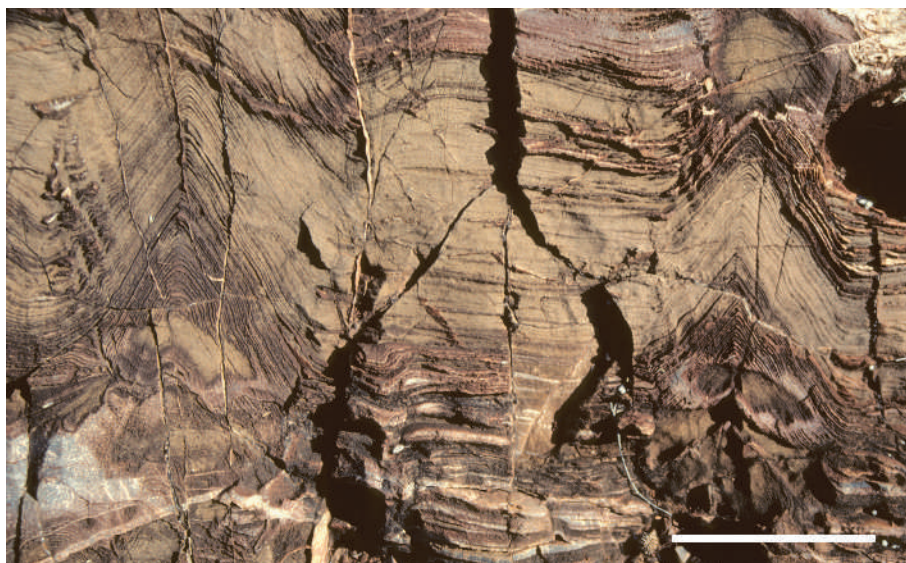


Figure 1 | Structure in the Strelley Pool Chert, Western Australia. These and many other forms are the subject of Allwood and colleagues' analyses², from which they conclude that microorganisms were adapted to and thriving in shallow marine environments some 3.4 billion years ago. Scale bar is 10 cm.

S. M. AWRAMIK

S. M. AWRAMIK



Figure 2 | Cutting through the complexities of stromatolites. Modern stromatolites — such as those shown here at Lee Stocking Island, Bahamas, as well as from places such as Shark Bay, Western Australia — provide essential comparisons for interpreting analogues from earlier in Earth's history.

of interspace areas: sediment in the interspace area has a vertical gradation in grain size, indicating the operation of mechanical processes, whereas grains in the laminae are not graded. Rather, these grains evidently accumulated on the steep slopes of the cones⁸, suggesting that there was some mechanism for trapping and binding them, like that found in microbially influenced sediment accretion. Allwood *et al.*² also point out that there are no known abiotic processes that could produce the facies-related stromatolites persistently and sometimes simultaneously on the carbonate platform.

No mechanical processes have been identified that form coniform stromatolites¹¹. But we do know of modern coniform stromatolites that are built by microorganisms. Growth experiments¹² on the microbes that build the modern cones indicate that the cone shape results from gliding by a photosynthetic filamentous microorganism (a cyanobacterium). Horizontally gliding filaments interfere with one another, form a clump, and other gliding filaments encounter the clump. These are deflected upwards towards the light and a cone results. Such structures could become fossilized by trapping and binding of sediment and/or by precipitation of mineral matter while the cones are growing. Coniform stromatolites make up about 44% of ten types of stromatolite recognized from 35 geological units ranging from 3.5 billion to 2.5 billion years in age¹³.

Hydrothermal environments have been favoured as the most likely places where life existed on the early Earth¹⁴. But the setting for the stromatolites of the Strelley Pool Chert indicates that microorganisms were adapted to and thriving in shallow marine environments, and that it is not necessary to invoke the presence of hydrothermal activity. The same conclusion applies to the microbial mats from the 3,416-million-year-old Buck Reef Chert in South Africa¹⁵.

So, microbial ecosystems evidently developed

in environments that we know (carbonate platforms) and produced structures that we recognize (stromatolites; Fig. 2). Why do we have to demand evidence that is more absolute? The prudent use of well-understood ancient analogues and modern examples to interpret stromatolites is a powerful scientific tool³. Actualism has served palaeontology well.

The late American comedian Rodney Dangerfield¹⁶ had a perpetual complaint: "I don't get no respect." The status of stromatolites as indicators of early life on Earth has suffered from a similar attitude. The work of Allwood *et al.*² will surely help to change that.

Stanley M. Awramik is in the Department of Earth Science, University of California, Santa Barbara, California 93106, USA.
e-mail: awramik@geol.ucsb.edu

1. Canfield, D. E. *Nature* **440**, 426–427 (2006).

2. Allwood, A. C., Walter, M. R., Kamber, B. S., Marshall, C. P. & Burch, I. W. *Nature* **441**, 714–718 (2006).
3. Riding, R. *Lethaia* **32**, 321–330 (1999).
4. Grotzinger, J. P. & Rothman, D. H. *Nature* **383**, 423–425 (1996).
5. Awramik, S. M. & Grey, K. *Proc. SPIE* **5906**, 59060P1–59060P9; doi:10.1117/12.625556 (2005).
6. Brasier, M., Green, O., Lindsay, J. & Steele, A. *Origins Life Evol. Biosphere* **34**, 257–269 (2004).
7. Lowe, D. R. *Nature* **284**, 441–443 (1980).
8. Hofmann, H. J., Grey, K., Hickman, A. H. & Thorpe, R. I. *Geol. Soc. Am. Bull.* **111**, 1256–1262 (1999).
9. Loucks, R. G. & Sarg, J. F. (eds) *Carbonate Sequence Stratigraphy* (Am. Assoc. Petrol. Geol., Mem. 57, 1993).
10. Grotzinger, J. P. & James, N. P. (eds) *Carbonate Sedimentation and Diagenesis in the Evolving Precambrian World* (Soc. Econ. Geol. Paleontol., Spec. Publ. 67, 2000).
11. Batchelor, M. T., Burne, R. V., Henry, B. I. & Slatyer, T. *Physica A* **350**, 6–11 (2005).
12. Walter, M. R., Bauld, J. & Brock, T. D. in *Stromatolites* (ed. Walter, M. R.) 273–310 (Elsevier, Amsterdam, 1976).
13. Hofmann, H. J. in *Microbial Sediments* (eds Riding, R. & Awramik, S. M.) 315–327 (Springer, Berlin, 2000).
14. Lindsay, J. F. *et al. Precamb. Res.* **143**, 1–22 (2005).
15. Tice, M. M. & Lowe, D. R. *Nature* **431**, 549–552 (2004).
16. http://en.wikipedia.org/wiki/Rodney_Dangerfield

PHOTONICS

Transparency on an optical chip

Robert W. Boyd and Daniel J. Gauthier

A two-laser trick that renders opaque media transparent can be achieved in systems of tiny optical resonators — with potentially profound consequences for optical communication and information processing.

The discovery of electromagnetically induced transparency (EIT) — an unusual effect that occurs when two laser beams interact within an optical material — and the use of novel techniques to fabricate ever smaller structures to control light have been recent exciting developments in optical physics. Writing in *Physical Review Letters*, Xu *et al.*¹ neatly combine the two, demonstrating an on-chip, all-optical analogue of EIT based on the response of coupled optical microresonators. The result may open up untrodden pathways in photonics, offering prospects of smaller, more efficient devices for the manipulation and transmission of light.

As originally implemented², EIT involves the interaction of laser light with a collection of atoms. It relies on the fact that when an incident photon has a 'resonant' energy equal to the difference in the energies of two levels of an atom, the photon can be absorbed by that atom and its energy used to excite the atom into the higher energy state. When two separate laser fields drive two such atomic transitions that share the same upper level (Fig. 1a, overleaf), destructive interference between the pathways connecting the upper level with the lower levels allows the quantum-mechanical probability that the atom is in the upper level to vanish. As there are no atoms in the upper level, there has been no absorption of the applied fields. The atom is thereby rendered 'transparent' to the applied laser fields in an

extremely narrow frequency range (Fig. 1b). This interference process has analogies in classical physics, where the coupling between two oscillators results in a reduction in the amplitude of their oscillation³.

The EIT phenomenon can be used to greatly strengthen nonlinear optical effects — such as the dependence of a material's refractive index on the intensity of the incident light — on which the operation of many photonic devices depend. In general, such nonlinear effects can be achieved by using as the source of the applied field a frequency-tunable laser tuned close to the resonance frequency of an atomic transition. Unfortunately, absorption of the laser light also increases at exactly these frequencies, and this effect undoes much of the benefit of working close to resonance. With two laser sources, however, quantum interference can be used to ensure that atomic absorption is eliminated, while retaining a large nonlinear response.

A downside of the 'traditional' atomic EIT effect — and more generally of nonlinear optics based on atomic resonance — is that it can be implemented only for light in a very small range of frequencies near fixed atomic transitions. An alternative way of achieving the transmission characteristics of EIT, as investigated by Xu and colleagues¹, uses small devices called optical microresonators⁴ (Fig. 1c), whose resonance frequency depends on their physical size. In such devices, resonance frequencies occur

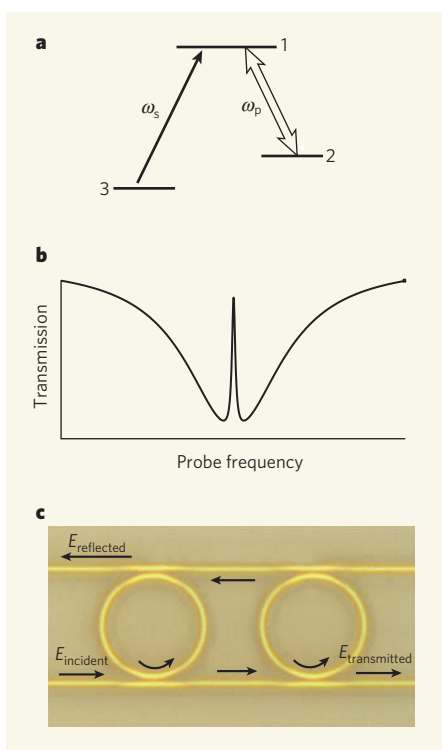


Figure 1 | Electromagnetically induced transparency. **a**, In conventional EIT in an atomic system, a strong pump field ω_p is tuned to an atomic energy transition (1–2) and creates a ‘transparency window’ for the signal field ω_s , which is tuned to a second transition (1–3). **b**, Destructive interference between the two absorption pathways results in a narrow spike of increased transmission within a wider absorption line. The refractive index of the material thus varies rapidly over a narrow frequency range, and the resulting increased dispersion can lead to ultraslow propagation of light. The narrow spike also provides a well-defined frequency marker for precision measurements. **c**, In a system of microresonators, studied by Xu *et al.*¹, EIT occurs because of destructive interference between two fields, leaving two appropriately spaced rings in the direction of the reflection port. The width of the absorption feature is determined by the overall loss of the ring resonator system (often dominated by coupling to the straight waveguide), whereas that of the transparency window is set by the internal loss of the ring resonators. (Part **c** adapted from ref. 1.)

whenever a whole number of wavelengths of the incident light fits into the resonator. Moreover, destructive interference occurs between light emerging from two coupled resonators.

Early microresonators were simply aerosols containing particles of varying sizes, but modern fabrication procedures produce individual resonators with highly controllable resonance frequencies⁴. These resonators come in various forms: ring shapes; disks or spheres, in which the light skims across the outer surface as sound does in a ‘whispering gallery’; or less obvious forms, such as defects in an otherwise perfect photonic crystal. Their small size makes them ideally suited to perform operations on light analogous to those performed by components

of silicon chips on electrons, and numerous applications have been proposed for them^{5–7}.

Several methods^{5,8,9} for combining EIT and optical microresonators to obtain EIT-like resonances in integrated optical systems were proposed in 2004, and a laboratory demonstration of such an effect was published last year¹⁰. In that experiment, EIT-like transmission spectra were observed in two interacting microresonators in the form of glass spheres approximately 400 micrometres in diameter that supported resonances of the whispering-gallery type.

Xu *et al.*¹ take this advance a stage further. Using modern nanofabrication procedures, they manufactured a pair of micro-ring resonators coupled to parallel waveguides on a silicon-on-insulator substrate typical of an electronic chip. The refractive index of the silicon that made up the waveguides and the rings (3.45) was significantly greater than that of the silicon oxide that surrounded them on all sides (1.46). This meant that light could be confined in rings of a very small diameter through the process known as total internal reflection. EIT-like behaviour was demonstrated by measuring the transmission at different frequencies of a laser beam injected into the system. The authors constructed several such devices, where the distance between the micro-rings was varied from device to device, and thus the spectral location of the EIT-like transmission peak could be adjusted.

So what good is all this? As the name indicates, EIT is primarily a way of rendering transparent an otherwise highly absorbing material, and so lends itself to applications such as the construction of long-distance optical transmission lines. The dispersion characteristic that results from EIT has already allowed ‘slow light’ propagation to be achieved^{11–13} at speeds just a small fraction of the normal speed of light. Optical ‘buffers’ to temporarily store light are an obvious

application of such slow-light technology.

Perhaps most importantly, EIT resonances can be much narrower than those of individual resonators, whether these are atoms or microresonators. As a resonance frequency can typically be measured only to an accuracy given by some fixed fraction of the resonance linewidth, narrowed linewidths could be useful for performing precise optical measurements of quantities such as magnetic field strength¹⁴. The ability to synthesize such components on an optical chip suggested by Xu and colleagues’ innovation¹ is a crucial step towards the development of such integrated optical devices.

Robert W. Boyd is at the Institute of Optics and Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA.

Daniel J. Gauthier is in the Department of Physics, Duke University, Durham, North Carolina 27708, USA.

e-mails: boyd@optics.rochester.edu; gauthier@phy.duke.edu

1. Xu, Q., Sandhu, S., Povinelli, M. L., Shukla, J., Fan, S. & Lipson, M. *Phys. Rev. Lett.* **96**, 123901 (2006).
2. Harris, S. E., Field, J. E. & Imamoglu, A. *Phys. Rev. Lett.* **64**, 1107–1110 (1990).
3. Alzar, C. L. G., Martinez, M. A. G. & Nussenzveig, P. *Am. J. Phys.* **70**, 37–41 (2002).
4. Vahala, K. J. *Nature* **424**, 839–846 (2003).
5. Yanik, M. F., Suh, W., Wang, Z. & Fan, S. *Phys. Rev. Lett.* **93**, 233903 (2004).
6. Scheuer, J. & Yariv, A. *Phys. Rev. Lett.* **96**, 053901 (2006).
7. Heebner, J. *et al.* *Opt. Lett.* **29**, 769–771 (2004).
8. Smith, D. D., Chang, H., Fuller, K. A., Rosenberger, A. T. & Boyd, R. W. *Phys. Rev. A* **69**, 063804 (2004).
9. Maleki, L., Matsko, A. B., Savchenkov, A. A. & Ilchenko, V. S. *Opt. Lett.* **29**, 626–628 (2004).
10. Naweed, A., Farca, G., Shopova, S. I. & Rosenberger, A. T. *Phys. Rev. A* **71**, 043804 (2005).
11. Hau, L. V., Harris, S. E., Dutton, Z. & Behroozi, C. H. *Nature* **397**, 594–598 (1999).
12. Boyd, R. W. & Gauthier, D. J. *Prog. Opt.* **43**, 497–530 (2002).
13. Bigelow, M. S., Lepeshkin, N. N. & Boyd, R. W. *Phys. Rev. Lett.* **90**, 113903 (2003).
14. Scully, M. O. & Fleischhauer, M. *Phys. Rev. Lett.* **69**, 1360–1363 (1992).

NEUROBIOLOGY

Extending influence

Eve Marder

Rather than merely firing in a digital on-off fashion, vertebrate neurons may have an analogue aspect to their signalling too — a finding that will not surprise many who have worked on invertebrate neurons.

Much of what we know about electrical signalling in the brain comes from extracellular recordings that detect when a neuron is firing action potentials. These recordings do not, however, provide continuous monitoring of the fluctuations of membrane potential, and do not capture sub-threshold changes in membrane potential such as those caused by individual synaptic events. The prevalence of

extracellular recordings in the literature has contributed to a collective consciousness in which the action potential or ‘spike’ is viewed as an invariant, all-or-nothing stereotyped event that occurs once a threshold membrane potential is reached. This ‘digital’ signal carries information from the neuronal cell body, the soma, down the axon to presynaptic terminals, where it evokes the release of neurotransmitter

to excite or inhibit the next neuron. In this simple framework, the spike frequency might influence the amount of transmitter release by mechanisms such as facilitation and depression¹, but sub-threshold events occurring in the soma and the dendrites (the projections that receive inputs from other neurons) are thought to be too far away to influence transmitter release from the axonal terminals.

Alle and Geiger writing in *Science*², and Shu *et al.* (this issue, page 761)³ now show, however, that the release of neurotransmitter from axon terminals of certain vertebrate neurons is influenced by the somatic membrane potential. So, in terms of neuronal signalling, the axon terminals and the soma are electrically much closer than many would have assumed (Fig. 1). The salient finding in both papers, from which all of the rest of the results follow, is that the length constant, λ , of axons is surprisingly long, at about 420–450 μm , in two types of neuron: hippocampal mossy fibres from rats² and layer 5 pyramidal cells from the prefrontal cortex of ferrets³. λ is the distance over which a voltage change imposed at one site will drop to approximately 37% of its initial value⁴. Although rapid changes in membrane potential are more attenuated by distance than are slow signals, when two regions of a neuron are less than λ apart, they are commonly considered to be electrically 'close'. Put another way, at a distance less than λ , changes in membrane potential at one place will appreciably alter the membrane potential at the other. In the case of the mossy fibres, λ was calculated using the distance between recordings made from boutons (presumably presynaptic release sites) and the soma². In the case of the prefrontal cortical neurons, simultaneous recordings were made from the soma and axon at known distances apart³.

How surprising is it that λ is so long in these neurons? Some readers will undoubtedly be taken aback to discover that vertebrate neurons can be so electrically compact, at least for slow signals. But those of us who work with large neurons in invertebrates already know of cases in which somatic voltages influence the axon more than a millimetre away⁵. In the vertebrate hippocampal² and layer 5 cortical neurons³ the consequence of λ being so long is that slow depolarizations (positive-going changes in membrane potential that bring the neuron closer to threshold) of the membrane at the somatic and dendritic regions are transmitted to the axon terminals, and can influence the release of transmitter.

Shu *et al.*³ show that depolarization of the soma can increase the amount of spike-mediated release of neurotransmitter from the presynaptic terminals by about 30% per 10 mV of somatic depolarization. Is this change likely to be important for the function of the circuits containing these neurons? This will depend on many factors, but it is fair to say that even small changes in synaptic strength can have significant influences on circuit performance.

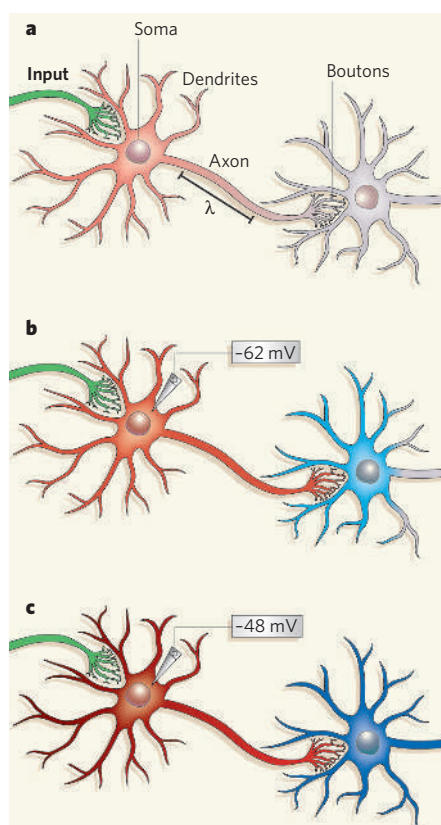


Figure 1 | Analogue signalling. The membrane potential of the soma can influence transmitter release from the axonal terminals^{2,3}. **a**, An input (green) produces a subthreshold depolarization of the neuron (pink). This depolarization spreads down the axon, decaying exponentially with distance. At distance λ , a voltage change in the soma will drop to 37% of its initial value. **b**, An action potential (red) is triggered from a baseline membrane potential of -62 mV . The active neuron releases neurotransmitter onto its follower neuron, resulting in a postsynaptic change (pale blue). **c**, The baseline somatic potential is depolarized to -48 mV , by barrages of input synapses or by an electrode. Now, when an action potential is evoked on the top of the baseline depolarization, the action potential (dark red) may be broader and results in a larger amplitude synaptic event in the follower neuron (dark blue).

In the work by Shu *et al.*³ somatic depolarization was associated with significant spike broadening, that is the return to resting potential occurred more slowly, in both the soma and the axon. Presumably this spike broadening may contribute to the enhanced transmitter release, as occurs elsewhere^{6,7}. In this study³, the authors recorded ongoing barrages of synaptic potentials resulting from spontaneous network activity. The depolarizations associated with these synaptic barrages were seen in both somatic and axonal recordings, demonstrating that signals of the amplitude and time-course of normal synaptic inputs could influence axonal membrane potential.

Alle and Geiger² saw no evidence of spike broadening in their recordings. But they did show that somatic depolarizations that mimic

those recorded *in vivo* during theta oscillations (an ongoing brain rhythm) can substantially increase the amplitudes of the postsynaptic currents evoked by action potentials.

Many types of invertebrate neuron release neurotransmitter both in response to action potentials and as a graded function of membrane potential^{8–11}. There is, however, an important difference between the results in the two studies highlighted here^{2,3} and the classic graded transmission well characterized in the retina and in many invertebrate neurons^{12–16}. In classic graded transmission, action potentials are not necessary for transmitter release because the threshold for transmitter release is often close to the resting potential. Therefore, synaptic release follows the analogue fluctuations of the presynaptic neuron's membrane potential. In contrast, the results of Alle and Geiger² and of Shu *et al.*³ show no indication that transmitter release occurs independently of action potentials, but only that the somatic membrane potential can influence how much transmitter is released by an action potential.

Over the years, I have watched the characteristics ascribed to vertebrate neurons slowly evolve to exhibit many of the attributes so clearly demonstrable in invertebrate neurons. I now find it gratifying that vertebrate neurons found in the hippocampus and cortex, brain regions well known for their importance in memory and cognition, may also avail themselves of some of the benefits of analogue processing in addition to spike-mediated release. My guess is that this work^{2,3} will trigger numerous studies to determine how transmitter release from spiking neurons in the vertebrate brain is influenced by slow changes in membrane potential that result from neuromodulation and ongoing activity. In this way, we will come to understand how neural circuits produce variable outputs as a function of behavioural state and mood.

Eve Marder is at the Volen Center and Biology Department, Brandeis University, 415 South St, Waltham, Massachusetts 02454-9110, USA. e-mail: marder@brandeis.edu

1. Zucker, R. S. & Regehr, W. G. *Annu. Rev. Physiol.* **64**, 355–405 (2002).
2. Alle, H. & Geiger, J. R. *Science* **311**, 1290–1293 (2006).
3. Shu, Y., Hasenstaub, A., Duque, A., Yu, Y. & McCormick, D. A. *Nature* **441**, 761–765 (2006).
4. Kandel, E. R., Schwartz, J. H. & Jessell, T. M. *Principles of Neural Science* (McGraw-Hill, New York, 2000).
5. Meyrand, P. *et al. J. Neurosci.* **12**, 2803–2812 (1992).
6. Sabatini, B. L. & Regehr, W. G. *J. Neurosci.* **17**, 3425–3435 (1997).
7. Ma, M. & Koester, J. *J. Neurosci.* **16**, 4089–4101 (1996).
8. Selverston, A. I. *et al. Prog. Neurobiol.* **7**, 215–290 (1976).
9. Ayali, A. *et al. J. Neurophysiol.* **79**, 2063–2069 (1998).
10. Angstadt, J. D. & Calabrese, R. L. *J. Neurosci.* **11**, 746–759 (1991).
11. Graubard, K. *et al. Proc. Natl Acad. Sci. USA* **77**, 3733–3735 (1980).
12. Graubard, K. *J. Neurophysiol.* **41**, 1014–1025 (1978).
13. Manor, Y. *et al. J. Neurosci.* **17**, 5610–5621 (1997).
14. DiCaprio, R. A. *J. Comp. Neurol.* **185**, 83–106 (1989).
15. Burrows, M. & Siegler, M. V. *J. Physiol. (Lond.)* **285**, 231–255 (1978).
16. Werblin, F. S. & Dowling, J. E. *J. Neurophysiol.* **32**, 339–355 (1969).

EARTHQUAKES

A hand on the aftershock trigger

Ian Main

Received wisdom relates that static stress change associated with an earthquake mainshock is the prime mover of its aftershocks. But a fresh look at the data points the finger at wave-surfing dynamic stress.

When a door is jammed, there are various things you can do to dislodge it — that is, to overcome the frictional resistance between it and its frame. You can nonchalantly push a bit harder (apply a static stress). Less elegantly, you can take a run at it and barge it open (apply a dynamic stress). And, if completely exasperated, you can rattle the handle back and forth (apply a cyclic stress). If you are already pushing very hard, it might take only a small extra heave or tap to free the door, leaving you making your entrance face down on the floor (feeling a large stress change from a small additional stress). Occasionally, tapping from the other side (applying a negative stress) can also help. Lastly, if you lean on the door long enough, it can give way spontaneously (through environmentally assisted weakening processes).

Such is the complexity of the physics of friction, even for a single pair of surfaces in contact. In this issue, Felzer and Brodsky (page 735)¹ try to pin down the relative contribution of some of these effects in a more complex context: the triggering of earthquake aftershocks. They examine the rate of decay with distance from the epicentre of the frequency of aftershocks following small and medium-sized earthquakes in the southern Californian catalogue. Their conclusion is that dynamic stresses radiated by seismic waves from the original shock dominate the aftershock-triggering mechanism over a wide range of distances between 0.2 and 50 kilometres from the fault rupture.

This comes as something of a surprise. From basic 'dislocation theory' for an earthquake source, one would expect the static stress changes induced by movement of rock in the original event to dominate aftershock generation near the fault. The effect of static stress should, compared with that of dynamic stress, decay quickly with distance: dynamic stress is transported over large distances by seismic waves.

If the orientations of both the original and a potential fault are known, dislocation theory can be used to calculate the changes induced by an earlier event in the static stress along, and at right angles to, the potential fault. By adding these up in a given proportion to define a resultant 'Coulomb' stress change, it is possible to assess whether these changes have brought a fault closer to failure². This method has been used in several case studies to predict the location both of aftershocks (defined as

being smaller than a mainshock) and more general triggered events (which can be the same size as, or very occasionally larger than, a triggering event).

Such studies often assume that triggered events are more likely when a fault is optimally oriented for failure from the start³. In real-life examples, the extra shove needed for triggering that emerges from these models often seems implausibly small: the largest aftershock of the Sumatra–Andaman earthquake of Boxing Day 2004, for instance, occurred where Coulomb stress was predicted to have increased by 10^4 Pa — just a tenth of atmospheric pressure⁴. Accordingly, 'rate and state friction'⁵, a phenomenon observed in the laboratory in which friction between rocks depends on slide velocity, has been invoked as a temporary stress-amplification effect³. An alternative hypothesis has it that the locations of triggered events are already primed to be in a near-critical state, making them susceptible to small stress perturbations². A combination of both effects is also a possibility.

But although the static triggering hypothesis works for earthquake pairs in many individual case studies, the method only marginally improves predictions of the location of large triggered events. A global study, based on a representative sample of earthquakes with known fault-plane orientations in the Centroid Moment Tensor (CMT) catalogue, found that only 61% of the triggered events occurred in areas of increased Coulomb stress⁶. Similarly, an analysis of all the CMT catalogue events showed no strong directional dependence of triggering frequency relative to the orientation of the potential mainshock fault planes⁷.

One suggestion is that static triggering works better when the effects of several — temporally distinct — large earthquakes are superimposed and short-range triggering effects (within one fault length) dominate. An example is a sequence of earthquakes that has worked along the North Anatolian fault in the past few hundred years, and which has focused attention on the next potential 'domino' in the line — in the Sea of Marmara near Istanbul⁸.

Felzer and Brodsky take a different tack¹. First, they look at aftershocks of small and medium-sized events, where the sampling statistics are good. Second, they ignore the direction of the aftershocks, and concentrate on how the frequency of triggered events declines

with increasing distance. They also restrict the maximum range considered to avoid contamination from background random seismicity, and, importantly, they use a new, extremely accurate catalogue of events⁹ with epicentres pinpointed to within some tens of metres.

The authors conclude that the frequency of aftershocks decays with distance as a single inverse power law with an exponent of around -1.35 . This decay is too gradual to be attributed to static stresses, which would in any case be implausibly tiny (less than 10 Pa) in the case of distant aftershocks of small events. But the exponent is consistent with the rate of attenuation of the maximum amplitude of seismic waves with distance when geometric spreading and wave absorption and scattering effects are taken into account. This implies that dynamic, rather than static, stress is the culprit.

Further work is required to pin down the exact mechanisms at work, however — notably the relative contributions to aftershock triggering of surface waves and shear waves (which travel through Earth's interior) emanating from a mainshock. Surface waves are in general more dispersed than shear waves, and, to return to the analogy of the door, apply more of an extended rattle than a sharp tap. They also decay in amplitude more slowly with distance, and their direction does not depend so strongly on the orientation of the original fault. Systematic variations in the frequency content and duration of the wave field could explain some of the scatter observed in the data obtained by Felzer and Brodsky¹. The nature of aftershocks at distances greater than 50 kilometres, where the data are dominated by background noise and are thus not considered by the authors, also needs attention. This will become possible only in the future, as more data on such distant aftershocks are required.

Dynamic stress transfer alone cannot explain the observed delay time for large events, or the observation of a power-law decay of the number of aftershocks with time — exactly the form predicted by the static processes of rate and state friction¹⁰ and environmentally assisted crack growth¹¹. Nevertheless, Felzer and Brodsky's work goes some way to satisfying us that earthquake aftershocks from small and moderate-sized original shocks, at least, are most likely to be triggered by a mechanism that is carried along on the crest of a wave. ■

Ian Main is in the School of Geosciences, University of Edinburgh, West Mains Road, Edinburgh EH9 3JW, UK.
e-mail: ian.main@ed.ac.uk

1. Felzer, K. R. & Brodsky, E. E. *Nature* **441**, 735–738 (2006).
2. Stein, R. S., King, G. C. P. & Lin, J. *Science* **265**, 1432–1435 (1994).
3. Stein, R. R. *Nature* **402**, 605–609 (1999).
4. Nalbant, S. S. *et al.* *Nature* **435**, 756–757 (2005).
5. Dieterich, J. H. *J. Geophys. Res.* **84**, 2161–2168 (1979).
6. Parsons, T. J. *Geophys. Res.* **107**, 2199;

- doi:10.1029/2001JB000646 (2002).
 7. Huc, M. & Main, I. G. *J. Geophys. Res.* **108**, 2324; doi:10.1029/2001JB001645 (2003).
 8. Hubert-Ferrari, A. *et al. Nature* **404**, 269–273 (2000).

9. Shearer, P., Hauksson, E. & Lin, G. *Bull. Seismol. Soc. Am.* **95**, 904–915 (2005).
 10. Parsons, T. A. *Geophys. Res. Lett.* **32**, L04302 (2005).
 11. Main, I. G. *Geophys. J. Int.* **142**, 151–161 (2000).

CHEMICAL TECHNOLOGY

All together now

Stephen J. Haswell

A method that combines techniques for performing chemical synthesis, separation and measurement on a single device illustrates the considerable potential of integrated lab-on-a-chip technology.

The traditional icons of chemistry, the test-tube and conical flask, are in many situations giving way to lab-on-a-chip methodology. This route to miniaturization offers high-throughput control of reactions that range in context from chemical synthesis and measurement, through biological processing, to medical diagnostics.

The prospect of realizing the full potential of lab-on-a-chip methodology comes a little closer with work reported by Detlev Belder and colleagues in *Angewandte Chemie*¹. The authors' strategy exploits the various benefits of performing reactions within a microfluidic environment, and interfaces this process to an efficient separation and detection system. They test their procedures by analysing the efficiency with which different mutant enzymes, created by directed evolution, select for different chiral or enantiomeric reaction products.

The potential advantages of lab-on-a-chip miniaturization include low operating volumes, faster reaction or sample-processing times, safer procedures, and the more efficient selective generation of products². For processes carried out in solution, many of these advantages arise from the control of samples and reagents that can be achieved within a network of micro-metre-sized fluidic channels. The temperature and concentration gradients created in such systems can be used to control reaction conditions; and the high ratios between surface area and volume can be especially effective in exploiting processes that occur at interfaces.

Microfluidic reactors do, however, have their drawbacks. For example, the price of trading the slow thermal and mass-transfer processes of a conventional reactor for the speed of a microreactor is much smaller amounts of product. But high flow rates or scaled-up procedures³ can compensate. For

instance, microreactors with channel cross-sections of only 250 μm have been used to carry out a common class of reaction, known as Mizoroki–Heck-type reactions, at a flow rate of 500 μl per minute to produce 12 g of pure product per hour⁴. So, in principle, 1,000 such reactors should be able to produce 288 kg of product per day. Another potential drawback is blockage of the channel network by

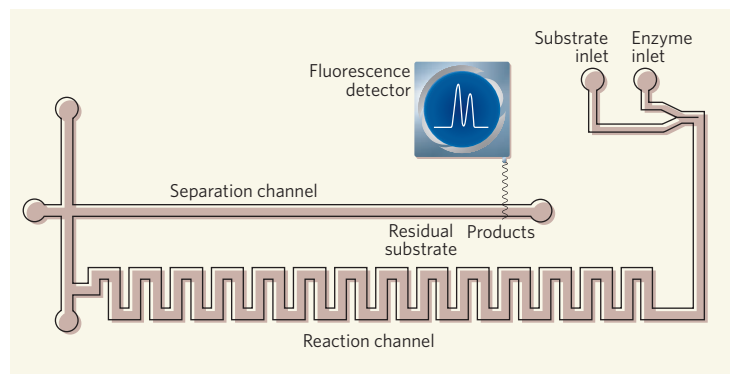


Figure 1 | Design and use of Belder and colleagues' microfluidic chip¹ for integrated chemical reaction and analysis. The system was tested on the enantioselectivity of fungal enzyme mutants. The substrate (glycidyl phenyl ether) and catalyst (epoxide hydrolase) mix in the meandering reaction channels. The reaction products and remaining substrate are separated electrophoretically in the separation channel, and the proportions of the product enantiomers are then measured with an external fluorescence detector. The chip is made of fused silica. It is about the size of a conventional microscope slide (7.5 × 2.5 cm) and a couple of millimetres thick. (Redrawn from ref. 1.)

particulate matter. But this problem can be minimized by careful selection of channel geometries and other parameters.

Belder and colleagues' chip for integrated catalysis and product analysis is shown in Figure 1. In their paper¹ they describe the reaction of 10 μl of glycidyl phenyl ether with a similar volume of various mutant epoxide hydrolases derived from the fungus *Aspergillus niger*. Enzyme and substrate were allowed to mix and react for 10 minutes, before an electrophoretic separation step (which took less than 90 seconds); fluorescence detection was used to monitor the proportions of different enantiomeric forms of product. To further illustrate the value of integrating the reaction and

separation–detection processes, the authors also used the chip to establish the reaction kinetics; this involved sequential injections of substrate and enzyme variants, followed by repeated analysis of the enantiomeric products as a function of time.

So far so good. But Belder and co-workers' chip¹ is also equipped with additional ports for introducing reactants, allowing up to three catalysts to be screened against one substrate in a single set of experiments. In principle, the current fabrication technology should allow this to be scaled up to tens or even hundreds of parallel streams. In terms of throughput, the authors also suggest that the 10-minute mixing and reaction time could be reduced through the introduction of a more efficient micromixer design. Finally, greater integration could be achieved by packaging the fluorescence-detection system, which at present requires an external laser source and detector to be aligned with the chip, into a single, more robust unit.

The contents of journals such as *Lab on a Chip* and *Microfluidics and Nanofluidics* attest to what has already been achieved with miniaturized chemistry. Building microreactors with full integration of chemical, biological, analytical and communication systems — for, say, remote operation — will require compromises, but it is an advantage that the fabrication of such systems has a common basis⁵. Uses of the technology will include drug discovery, where drug candidates could be synthesized and screened for toxicity and activity in minutes, with an iterative procedure allowing swifter and surer identification of the best molecules⁶. There will be other applications in personalized medicine, drug delivery, environmental and security monitoring, and forensics and space exploration.

Finally, integrated miniaturized systems should eventually be relatively cheap. They will be much more affordable than full lab facilities — not least in developing countries.

Stephen J. Haswell is in the Department of Chemistry, University of Hull, Cottingham Road, Kingston upon Hull HU6 7RX, UK.
 e-mail: s.j.haswell@hull.ac.uk

- Belder, D., Ludwig, M., Wang, L.-W. & Reetz, M. T. *Angew. Chem. Int. Edn* **45**, 2463–2466 (2006).
- Watts, P. & Haswell, S. J. *Chem. Soc. Rev.* **34**, 235–246 (2005).
- Jahnisch, K., Hessel, V., Lowe, H. & Baerns, M. T. *Angew. Chem. Int. Edn* **43**, 406–446 (2004).
- Liu, S., Fukuyama, T., Sato, M. & Ryu, I. *Org. Process Res. Dev.* **8**, 477–481 (2004).
- Vilkner, T. *et al. Anal. Chem.* **76**, 3373–3386 (2004).
- Watts, P. & Haswell, S. J. *Drug Discov. Today* **8**, 586–593 (2003).

BRIEF COMMUNICATIONS

Rediscovery of the world's leggiest animal

This ancient animal, found only in a tiny stretch of California, is close to being a true millipede.

The millipede species *Illacme plenipes* comes the closest to having its namesake's mythical 1,000 legs — individuals can bear up to 750 legs¹. Here we record the rediscovery of this extremely rare species, which has not been reported since its original description¹ some 80 years ago, at a tiny locality of 0.8 km² in San Benito County, California. Because of the rarity and narrow geographical range of this delicate species, its fragile habitat must be protected at all costs.

The arthropod class Diplopoda (millipedes) contains some 10,000 described species² and is one of the most ancient groups of terrestrial animals^{3–5}. Confident reconstructions of diplopod phylogeny and the taxon's placement within the broader context of arthropod evolution remain elusive, however, due in part to the group's antiquity, but also to the rarity of key taxa^{6,7} — some are known from just a single specimen. The millipede *I. plenipes* (Arthropoda: Myriapoda: Diplopoda: Siphonurhinidae) has not been sighted since 1926 (ref. 1).

Over the course of three separate collecting expeditions, we captured four male, three female and five juvenile specimens of *I. plenipes*. The female specimens were 32.38–33.20 mm long and comprised 170 or 171 segments with a total of 662–666 legs — slightly fewer than the record-holding specimen first described¹. Male specimens were smaller at 14.38–16.15 mm, with 84–105 segments and 318–402 legs.

Millipedes undergo anamorphosis, the addition of segments and legs during post-embryonic growth⁸. In siphonurhinid millipedes, this continues for an indeterminate length of time, possibly even after sexual maturity. This would result in continuous segment addition and elongation and would account for the length variation we observed in our adult specimens. The average body width for all specimens is a thread-like 0.57 mm.

The species' simple outward appearance (Fig. 1a) belies its ornate exoskeletal surface structure, which becomes evident through scanning electron microscopy (Fig. 1b–e). (For

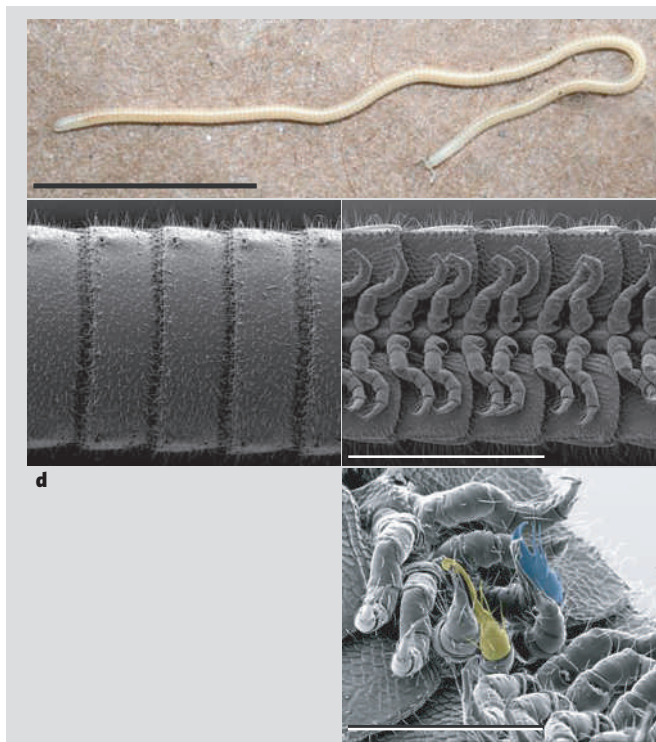


Figure 1 | The millipede *Illacme plenipes*. **a**, Live female *I. plenipes* with 662 legs (captured by Rob Marek). **b, c**, Scanning electron micrographs of male *I. plenipes*: **b**, mid-section, dorsal aspect of segments (note dense setation and micro-sculpturation of intersegmental cuticle); **c**, mid-section, ventral aspect of segments showing paired appendages on each diplo-segment. **d**, Close-up view of ozopore, the opening to the repugnatorial gland, and surface structure of cuticular projection on posterior margin of segment. **e**, Male copulatory device (gonopods), which are modified legs (ninth and tenth pair); left anterior is shaded blue; right posterior is in yellow. Scale bars: **a**, 10 mm; **b, c**, 400 μ m; **d**, 50 μ m; **e**, 200 μ m.

more images and for the only live video ever recorded, to our knowledge, of *I. plenipes*, see supplementary information.) Each segment is equipped with numerous hairs (setae) and previously undocumented cuticular projections. Many of the setae on the dorsal segmental plates (tergites) seem to secrete a 'silk-like' substance whose function is unknown. The trailing edge of each tergite and the ozopore — a gland opening — have bizarre, almost gothic ornamentation.

The male copulatory device (Fig. 1e) is far more complicated than previously described⁹, comprising an anterior component that cups the posterior articles *in situ*. The distal article of the posterior gonopod, the sixth podomere, is bifurcate and each branch is armed with a serrated apex like a bird claw

(Fig. 1e, structure shaded yellow).

The rediscovery of this remarkable animal in a known biodiversity hotspot, the California Floristic Province^{10,11}, is noteworthy: other species of closely related¹² and equally rare siphonurhinid genera are endemic to centres of biodiversity in Indo-Burma, southern Africa, Sundaland and Wallacea. The family's enigmatic distribution further supports the unique status of these hotspots as repositories of exceptional diversity and provides evidence that the siphonurhinid radiation pre-dates the break-up of Pangaea more than 200 million years ago.

Our rediscovery of *I. plenipes* at a time when sophisticated microscopes are available has revealed the fine-scale structure of a creature whose morphology is surprisingly intricate. The Diplopoda are an intriguing class of organisms that await new discoveries.

Paul E. Marek, Jason E. Bond
Department of Biology, East Carolina University, Greenville, North Carolina 27858, USA
e-mail: pm0623@ecu.edu

1. Cook, O. F. & Loomis, H. F. *Proc. US Natl Mus.* **72**, 1–26 (1928).
2. Hoffman, R. L., Golovatch, S., Adis, J. & de Morais, J. in *Amazonian Arachnida and Myriapoda* (ed. Adis, J.) 505–533 (Pensoft, 2002).
3. Wilson, H. M. & Anderson, L. I. *J. Paleontol.* **78**, 169–184 (2004).
4. Edgecombe, G. D. *Nature* **394**, 172–175 (1998).
5. Regier, J. C., Wilson, H. M. & Shultz, J. W. *Mol. Phylogenet. Evol.* **34**, 147–158 (2005).
6. Sierwald, P., Shear, W. A., Shelley, R. M. & Bond, J. E. *J. Zool. Syst. Evol. Res.* **41**, 87–99 (2003).
7. Shear, W. A. & Leonard, W. P. *Zootaxa* **243**, 1–11 (2003).
8. Enghoff, H., Dohle, W. & Blower, J. G. *Zool. J. Linn. Soc.* **109**, 103–234 (1993).
9. Shelley, R. M. *Myriapodologica* **4**, 21–33 (1996).
10. Myers, N. *Environmentalist* **8**, 187–208 (1988).
11. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. *Nature* **403**, 853–858 (2000).
12. Hoffman, R. L. *Classification of the Diplopoda* (Muséum d'Histoire Naturelle, Geneva, 1979).

Supplementary information accompanies this communication on Nature's website.
Received 2 March; accepted 11 May 2006.
Competing financial interests: declared none.
doi:10.1038/441707a

BRIEF COMMUNICATIONS ARISING online
www.nature.com/bca see Nature contents.

MOLECULAR SPECTROSCOPY

Complexity of excited-state dynamics in DNA

Arising from: C. E. Crespo-Hernández, B. Cohen & B. Kohler *Nature* **436**, 1141–1144 (2005).

Absorption of ultraviolet light by DNA is known to lead to carcinogenic mutations, but the processes between photon absorption and the photochemical reactions are poorly understood. In their study of the excited-state dynamics of model DNA helices using femtosecond transient absorption spectroscopy¹, Crespo-Hernández *et al.* observe that the picosecond component of the transient signals recorded for the adenine–thymine oligonucleotide (dA)₁₈•(dT)₁₈ is close to that for (dA)₁₈, but quite different from that for (dAdT)₉•(dAdT)₉; from this observation, they conclude that excimer formation limits excitation energy to one strand at a time. Here we use time-resolved fluorescence spectroscopy to probe the excited-state dynamics, which reveals the complexity of these systems and indicates that the interpretation of Crespo-Hernández *et al.*

is an oversimplification. We also comment on the pertinence of separating base stacking and base pairing in excited-state dynamics of double helices and question the authors' assignment of the long-lived signal component found for (dA)₁₈•(dT)₁₈ to adenine excimers.

Figure 1a shows the fluorescence decays of (dA)₂₀ at three different wavelengths. Combining these results with our previous measurements obtained for (dA)₂₀ by femtosecond fluorescence upconversion², at least five exponentials are needed to fit the decays over the 100 femtoseconds to 20 nanoseconds time range. A crucial point is that all time constants vary strongly with the emission wavelength. The same effect is encountered for (dAdT)₁₀•(dAdT)₁₀ and has been reported previously³ for poly(dA)•poly(dT).

We interpret this complex behaviour in terms of a model in which a large number of excited states are formed that are delocalized over several bases, which may be located both on the same strand and on opposite strands, with subsequent energy transfer³. This model is based on calculations made in the frame of exciton theory and combines quantum chemistry data and molecular dynamics simulations^{4–7}, and accounts not only for the fluorescence decays but also the steady-state absorption and fluorescence spectra. Delocalization of the excitation energy is governed by the electronic coupling, which depends on the oligomer's conformation. Conformational changes occurring on pico- and nanosecond timescales are controlled by an ensemble of interactions involving not only the bases but also the backbone, counterions and water molecules^{5,8}. In this sense, both base stacking and base pairing determine excited-state dynamics.

For these reasons, the time constants provide a phenomenological description of the decays and do not correspond to specific excited states. However, it is possible to make a rough comparison of the overall excited-state dynamics by considering the decays recorded at the maxima of the fluorescence spectra of the three oligomers (Fig. 1b). In the case of (dAdT)₁₀•(dAdT)₁₀ and (dA)₂₀, the fluorescence maxima have been assigned to excimer emission^{9,10}. We observe that, in contrast to the transient absorption signals, the fluorescence decay obtained for (dA)₂₀•(dT)₂₀ on the sub-nanosecond timescale is much shorter than that for (dA)₂₀. The same observation is valid when comparing the decays of these single and double strands at identical wavelengths.

Neither in fluorescence nor in transient absorption experiments is the amplitude of the detected signals proportional to the excited-

state population. The transient absorption signals depend on the difference between the molar extinction coefficients of the $S_0 \rightarrow S_1$ and $S_1 \rightarrow S_n$ transitions (where S_0 and S_1 denote the ground and first excited electronic states and S_n denotes different, higher excited electronic states) at the probed wavelengths. As the steady-state absorption spectra of these oligomers correspond to a large number of transitions⁶ and nothing is known about the $S_1 \rightarrow S_n$ spectra of their various excited states, the percentage of the 'excimer' population reported by Crespo-Hernández *et al.* is not necessarily correct.

The important difference between the transient absorption and fluorescence decays of (dA)_n•(dT)_n indicates the formation of dark transient species, as also discussed by Crespo-Hernández and colleagues^{1,10}. If these dark species are adenine 'excimers', they must have a different electronic structure from the fluorescent 'excimers' of (dA)_n, and, therefore, different lifetimes. Consequently, the similar time constants observed by transient absorption for (dA)₁₈•(dT)₁₈ and (dA)₁₈ may be fortuitous. The species observed by transient absorption in (dA)₁₈•(dT)₁₈ could as well be interstrand A–T charge-transfer states, as suggested by theoretical calculations¹¹. The behaviour of such states in deuterated water, as examined by Crespo-Hernández *et al.*, is not readily predictable because water molecules form a variety of interstrand and intrastrand bridges between bases¹².

Dimitra Markovitsi, Francis Talbot, Thomas Gustavsson, Delphine Onidas, Elodie Lazzarotto, Sylvie Marguet

Laboratoire Francis Perrin, CEA/DSM/DRECAM/SPAM-CNRS URA 2453, CEA Saclay, 91191 Gif-sur-Yvette, France
e-mail: dimitra.markovitsi@cea.fr

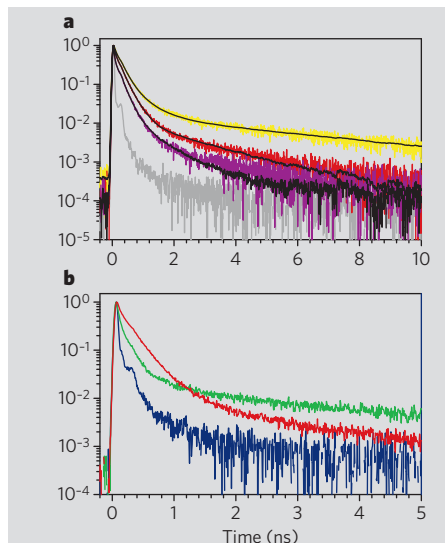


Figure 1 | Fluorescence decays recorded by time-correlated single-photon counting. **a**, (dA)₂₀ at 330 nm (magenta), 360 nm (red) and 420 nm (yellow). Black lines correspond to fits with multi-exponential functions, yielding the following sets of time constants (3.2, 37.5, 186 and 748 ps), (11.6, 101, 253 and 1,830 ps) and (39, 198, 551 and 5,050 ps), respectively. The instrumental response function (grey) is represented by the sub-picosecond fluorescence decay at 330 nm of thymidine-5'-monophosphate. **b**, (dA)₂₀ at 330 nm (red), (dA)₂₀•(dT)₂₀ at 330 nm (blue), (dAdT)₁₀•(dAdT)₁₀ at 420 nm (green). **Methods.** DNA oligomers (from Eurogentec) dissolved in phosphate buffer (pH 6.8) were excited by femtosecond pulses (100 fs, 267 nm). All decay curves correspond to the total fluorescence $F(t)$ constructed from the parallel (I_{par}) and perpendicular (I_{perp}) components according to: $F(t) = I_{\text{par}}(t) + 2GI_{\text{perp}}(t)$, where G accounts for the polarization-dependent sensitivity of the detection system. For further experimental details, see ref. 3.

- Crespo-Hernández, C. E., Cohen, B. & Kohler, B. *Nature* **436**, 1141–1144 (2005).
- Markovitsi, D., Sharonov, A., Onidas, D. & Gustavsson, T. *Chem. Phys. Chem.* **3**, 303–305 (2003).
- Markovitsi, D., Onidas, D., Gustavsson, T., Talbot, F. & Lazzarotto, E. *J. Am. Chem. Soc.* **127**, 17130–17131 (2005).
- Bouvier, B., Gustavsson, T., Markovitsi, D. & Millié, P. *Chem. Phys.* **275**, 75–92 (2002).
- Bouvier, B. *et al.* *J. Phys. Chem. B* **107**, 13512–13522 (2003).
- Emanuele, E., Markovitsi, D., Millié, P. & Zakrzewska, K. *Chem. Phys. Chem.* **6**, 1387–1392 (2005).
- Emanuele, E., Zakrzewska, K., Markovitsi, D., Lavery, R. & Millié, P. *J. Phys. Chem. B* **109**, 16109–16118 (2005).
- Hartmann, B. & Lavery, R. *Q. Rev. Biophys.* **29**, 309–368 (1996).
- Ge, G. & Georgeiou, S. *Photochem. Photobiol.* **54**, 301–305 (1991).
- Crespo-Hernández, C. E. & Kohler, B. *J. Phys. Chem. B* **108**, 11182–11188 (2004).
- Starikov, E. B., Lewis, J. P. & Sankey, O. F. *Int. J. Mod. Phys. B* **19**, 4331–4357 (2005).
- Kopka, M. L., Fratini, A. V., Drew, H. R. & Dickerson, R. E. *J. Mol. Biol.* **163**, 129–146 (1983).

doi:10.1038/nature04903

MOLECULAR SPECTROSCOPY

Crespo-Hernández et al.

Replying to: D. Markovitsi et al. *Nature* **442**, doi:10.1038/nature04903 (2006)

We have shown¹ that long-lived excited electronic states known as excimers², which arise from base stacking³, are formed in high yields in a variety of synthetic DNA oligonucleotides. Markovitsi et al.⁴ question our interpretation, and claim that these states can be accounted for by their exciton theory. However, neither this nor their emission data contradict our finding that in single- and double-stranded A•T (adenine–thymine-paired) DNA, excited states decay through long-lived intermediate states in which excitation is shared by stacked bases.

Time-resolved fluorescence and transient absorption signals from nucleic acids can differ in their dynamics^{1,3} and result from different populations. Emission experiments are sensitive to states with significant radiative transition probability, whereas transient absorption can detect states that are dark in emission, such as the long-lived states in DNA^{1,3}. Fluorescence provides an important but narrow view of the decay pathways taken by excess electronic energy in DNA. Less than 0.1% of all excitations in DNA decay by photon emission² and transient absorption¹, and various time-resolved ionization spectroscopies^{5–7} provide information about decay pathways that involve dark states.

The different fluorescence decays of Markovitsi et al. for (dA)₁₈•(dT)₁₈ versus (dA)₁₈ disagree with a time-resolved emission study in which long-lived decays were identical for analogous 15-mer oligonucleotides⁸. Both studies had similar time resolution, which is adequate for characterizing the components with lifetimes of more than 100 picoseconds that dominate the transient absorption signals. It is difficult to address this discrepancy as

Markovitsi et al. compare signals at different wavelengths in their Fig. 1b, and provide lifetimes for (dA)₂₀ but not for (dA)₂₀•(dT)₂₀. They also report faster emission decays for (dA)₂₀•(dT)₂₀ than for (dA)₂₀ on the sub-nanosecond and nanosecond timescale (their Fig. 1b), but the reverse ordering is reported by the same group on the picosecond timescale (Fig. 3 of ref. 9). These uncertainties must be resolved before further discussion is warranted.

Markovitsi et al.⁴ claim that modelling of DNA (their refs 4–7) accounts for decays, steady-state absorption and fluorescence spectra, where exciton states are calculated from dipolar couplings among the ¹ππ* states of each constituent base. Although a first step towards modelling steady-state absorption spectra, none of the studies uses this theory to account for trends in steady-state emission spectra or lifetimes, as suggested⁴. The theory omits orbital overlap interactions, which are important for chromophores separated by small distances¹⁰, and cannot describe the charge-transfer states that arise in DNA^{1,11,12}. In our paradigm¹, charge-transfer states are the intermediates that the initial bright states pass through to reach the electronic ground state.

No evidence is provided by Markovitsi et al. that their emission signals represent the principal decay pathway. In contrast, our transient absorption signals at 250 nm indicate that most excitations in stacked bases decay to long-lived states¹. We dispute their argument that the amplitude of fluorescence and transient absorption signals are not proportional to excited-state populations: such signals are proportional to a sum over the population of each state present, multiplied by its transition

cross-section. Target analysis can be used¹³ to recover time-dependent populations and yields from these signals with a simple kinetic model. Bleach recovery directly probes ground-state repopulation, and the sum of amplitudes of the slower exponentials provides a yield of the long-lived excimer states: if excimers absorb at 250 nm, then this yield is a lower bound¹. In any case, similar transient absorption decays are observed for (dA)₁₈ and (dA)₁₈•(dT)₁₈ (Fig. 4a of ref. 1). The same long-lived states seen in the adenine single strand are seen in the duplex, indicating that base pairing does not introduce ultrafast quenching pathways. This is why vertical base stacking, and not base pairing, plays the primary role in DNA excited-state dynamics.

Carlos E. Crespo-Hernández, Boiko Cohen, Bern Kohler

Department of Chemistry, The Ohio State University, Columbus, Ohio 43210, USA
e-mail: kohler@chemistry.ohio-state.edu

1. Crespo-Hernández, C. E., Cohen, B. & Kohler, B. *Nature* **436**, 1141–1144 (2005).
2. Crespo-Hernández, C. E., Cohen, B., Hare, P. M. & Kohler, B. *Chem. Rev.* **104**, 1977–2019 (2004).
3. Crespo-Hernández, C. E. & Kohler, B. *J. Phys. Chem. B* **108**, 11182–11188 (2004).
4. Markovitsi, D. et al. *Nature* **442**, doi:10.1038/nature04903 (2006).
5. Kang, H., Jung, B. & Kim, S. K. *J. Chem. Phys.* **118**, 6717–6719 (2003).
6. He, Y., Wu, C. & Kong, W. *J. Phys. Chem. A* **107**, 5145–5148 (2003).
7. Ullrich, S., Schultz, T., Zgierski, M. Z. & Stolow, A. *J. Am. Chem. Soc.* **126**, 2262–2263 (2004).
8. Plessow, R., Brockhinke, A., Eimer, W. & Kohse-Hoeinghaus, K. *J. Phys. Chem. B* **104**, 3695–3704 (2000).
9. Markovitsi, D., Sharonov, A., Onidas, D. & Gustavsson, T. *Chem. Phys. Chem.* **3**, 303–305 (2003).
10. Scholes, G. D. & Ghiggino, K. P. *J. Phys. Chem.* **98**, 4580–4590 (1994).
11. Wilson, R. W. & Callis, P. R. *J. Phys. Chem.* **80**, 2280–2288 (1976).
12. Starikov, E. B. *Mod. Phys. Lett. B* **18**, 825–831 (2004).
13. van Stokkum, I. H. M., Larsen, D. S. & van Grondelle, R. *Biochim. Biophys. Acta Bioenerget.* **1657**, 82–104 (2004).

doi:10.1038/nature04904

Titan Radar Mapper observations from Cassini's T₃ fly-by

C. Elachi¹, S. Wall¹, M. Janssen¹, E. Stofan², R. Lopes¹, R. Kirk³, R. Lorenz^{4,5}, F. Paganelli¹, L. Soderblom³, C. Wood⁶, L. Wye⁷, H. Zebker⁷, Y. Anderson¹, S. Ostro¹, M. Allison⁸, R. Boehmer¹, P. Callahan¹, P. Encrenaz⁹, E. Flamini¹⁰, G. Francescetti¹¹, Y. Gim¹, G. Hamilton¹, S. Hensley¹, W. Johnson¹, K. Kelleher¹, D. Muhleman¹², G. Picardi¹³, F. Posa¹⁴, L. Roth¹, R. Seu¹³, S. Shaffer¹, B. Stiles¹, S. Vetrella¹¹ & R. West¹

Cassini's Titan Radar Mapper imaged the surface of Saturn's moon Titan on its February 2005 fly-by (denoted T₃), collecting high-resolution synthetic-aperture radar and larger-scale radiometry and scatterometry data. These data provide the first definitive identification of impact craters on the surface of Titan, networks of fluvial channels and surficial dark streaks that may be longitudinal dunes. Here we describe this great diversity of landforms. We conclude that much of the surface thus far imaged by radar of the haze-shrouded Titan is very young, with persistent geologic activity.

The Cassini Titan Radar Mapper^{1,2} is a four-mode, K_u-band (wavelength = 2.17 cm) instrument. The synthetic-aperture radar (SAR) mode produces image data with best resolution of 300 m per pixel; scatterometry, altimetry and radiometry modes produce lower-resolution data. Here we report on data acquired during the T₃ fly-by and compare these with the T_A data, acquired in October 2004 at higher latitudes in the same hemisphere³.

The T₃ encounter, with a closest approach at 30° N latitude, 70° W longitude, and 1,579 km altitude at 6:58 Universal Time Coordinated (UTC) on 15 February 2005, was Cassini's third fly-by of Titan and the first after the Huygens probe descent. Radar observations started with an inbound altimetry track covering an arc of ~8°, followed immediately by a 40-min SAR swath, symmetric about closest approach (Fig. 1a; an uncontrolled, higher-resolution version is in the Supplementary Information). The SAR imaging data covered approximately 1.8 × 10⁶ km², with a spacecraft-to-surface range up to 5,500 km and a spatial resolution of 300–1,500 m. At 20 min after closest approach, an outbound altimetry track began, after which scatterometry scans extended from about 35° S to 50° N and 0° to 15° W, overlapping and extending similar coverage from T_A. Low-resolution radiometry was obtained concurrently with all other data types.

Mapping units

In the T_A SAR swath, five surface units were mapped, based solely on brightness variations, general plan-form shape, and texture³. These are not classified as material stratigraphic units (such as rock or ice) because we do not fully understand the relative contributions to SAR brightness from surface roughness, surface topography, material composition, and/or volume scattering. The T₃ swath contains all five of these units (dark homogeneous, mottled, dark mottled, bright lobate, and bright lineated) and three additional units (bright mottled, bright homogeneous, and bright rough) (see Fig. 2). Mapping in the T₃ swath is complicated by long and narrow dark

features, occurring on all mapped units—except dark mottled and bright lobate—that are likely to be surficial deposits obscuring the underlying units.

The dominant unit in T₃, covering almost half of the swath and characterized by irregular bright patches of variable size, is a bright mottled unit (Fig. 2a and b) not seen in the T_A swath. Small (<10 km) bright spots in this unit may be topographic features. The second most extensive is the mottled unit (Fig. 2c), occurring in the eastern half of the swath. It includes gradational boundaries, moderate to bright backscatter, and outcroppings of the dark mottled unit. The bright homogeneous unit (Fig. 2d) occurs in the centre of the T₃ swath and contains a large impact crater and numerous channel-like features that we interpret as fluvial in origin. The low-backscatter and relatively featureless dark homogeneous unit (Fig. 2e), the most extensive in the T_A swath, occurs in the eastern portion of the T₃ swath.

We identified a bright lobate unit in the T_A swath and interpreted it as cryovolcanic in origin on the basis of its resemblance in shape and brightness to SAR images of basaltic flows on Earth³. Fluvial or mass-wasting mechanisms are less likely possibilities, but the previous identification of two dome-like features^{3,4} and theoretical considerations^{4,5} also indicate that cryovolcanism probably happened on Titan. Three small features in the T₃ swath are mapped as the same bright lobate unit (Fig. 2e), though the relatively low spatial resolution in relation to the size of the features hinders interpretation. In T₃ these patches of bright lobate unit appear in association with two bright hills and a circular feature. The T_A bright lineated unit is an SAR-right, lineated unit with polygonal boundaries. It occurs in one location in the T₃ swath (Fig. 2d).

Our interpretation is that these were formed by a flowing material, in places confined by topography that may be tectonic in origin, because of its combination of fan-shaped and sinuous features. If this unit is a flow deposit it is probably composed of rough ice materials. The lineations are either produced by flow or are dark streaks

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA. ²Proxemy Research, Bowie, Maryland 20715, USA. ³US Geological Survey, Flagstaff, Arizona 86001, USA. ⁴Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA. ⁵IFSI-INAF, Via del Fosso del Cavaliere 100, 00133 Rome, Italy. ⁶Planetary Science Institute, Tucson, Arizona 85719, USA. ⁷Stanford University, Stanford, California 94305, USA. ⁸Goddard Institute for Space Studies, National Aeronautics and Space Administration New York, New York 10025, USA. ⁹Observatoire de Paris, 92195 Meudon, France. ¹⁰Alenia Aerospazio, 00131 Rome, Italy. ¹¹Facoltà di Ingegneria, 80125 Naples, Italy. ¹²Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA. ¹³Università La Sapienza, 00184 Rome, Italy. ¹⁴INFN and Dipartimento Interateneo di Fisica, Politecnico di Bari, 70126 Bari, Italy.

overlying the brighter underlying deposit. The bright lined unit was also mapped in the T_A swath, both in association with fan-shaped and sinuous features and as an angular unit associated with dark terrain. The dark mottled unit, occurring in crescent-shaped patches in the T_A swath, was interpreted as smooth deposits, perhaps ponds of hydrocarbons³. Similar terrain, forming irregularly shaped patches tens of kilometres across, occurs in the eastern end of T_3 (Fig. 2c). No surficial streaks are seen on or associated with these patches, suggesting that the T_3 dark mottled unit may be liquids or smooth patches of particulates (ice sand) undisturbed by wind.

In the eastern part of the T_3 swath we have identified a bright rough unit (Fig. 2e), which consists of groups of materials that appear to be topographically high and rugged, as they exhibit evidence of both layover and topographic shading. These features are more distinct and noteworthy than the small hills found in the T_A unit. We interpret this material to be high-standing, fractured ice, older than the surrounding plains material that appears to embay it.

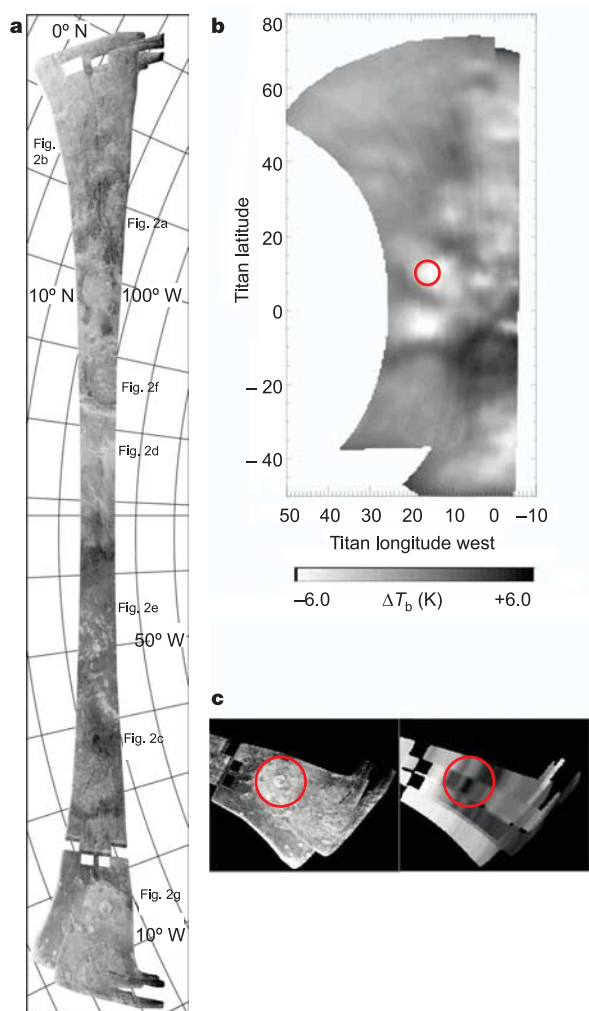


Figure 1 | T_3 SAR swath. **a**, Full swath, indicating locations of enlargements shown in Fig. 2b. **b**, Relative brightness temperature ΔT_b accompanying the scatterometry in Fig. 3 for T_A outbound and T_3 passes. The red circle indicates a radiometrically cool feature surrounding the 80-km crater Sinlap, a feature shown at higher resolution in **c**. **c**, The left panel shows the end of the SAR strip with Sinlap and surrounding terrain, interpreted as an ejecta blanket. The right panel shows the accompanying radiometry. The red circle defines the same area as in **b**. Spatial resolution is determined by the footprints of the five SAR beams spanning the strip.

Impact craters

No impact craters were positively identified in the T_A SAR image swath³. The T_3 SAR swath is dominated by a 450-km-wide circular structure that is clearly an impact basin. A second, 80-km-diameter impact crater appears at the eastern end of T_3 and has been subsequently observed at infrared wavelengths^{4,6}. Both craters occur in the bright homogeneous unit and on the boundary of the mottled and bright mottled units.

The 450-km-diameter double-ring impact basin, Menrva, is centred near 87° W, 19° N (Fig. 2f); its general shape is also seen in the near-infrared image shown as background⁶. The steep inner wall is bright, exhibiting radial lineaments, presumably grooves and chutes. The southern and western regions of the floor are relatively bland, suggesting resurfacing. The centre of the basin appears elevated and rough-textured, with a bright inner ring about 100 km in diameter. Dark, thin, linear streaks extend from the basin's lower wall onto the basin floor. The western rim shows more signs of erosion than the eastern rim. Candidate fluvial features, discussed below, may be associated with the basin.

The 80-km crater Sinlap, at 16° W, 11° N, shows no evidence of a raised rim (Fig. 2g). Craters of this diameter on Ganymede⁷ have domed floors, probably due to viscous relaxation, and central pits. There is no indication of such features in Sinlap, nor of a central peak or peak ring. A 2- to 3-km bright spot appears 8 km west-southwest of the centre. The floor seems flat, similar to lava-flooded craters on the Moon or to craters with smooth deposits on Mars. Assuming that the crater wall has the same height and slope around its perimeter, we calculate⁸ a wall slope of $16 \pm 5^\circ$ and a crater depth of $1,300 \pm 200$ m for a depth/diameter ratio of 0.016 ± 0.03 . The crater is asymmetrically surrounded by a blanket of SAR-bright material, presumably impact ejecta, biased towards the eastern side. In places it extends more than two crater radii beyond the rim, possibly arguing for some degree of fluidization of the ejecta⁹. No extensive distal ejecta deposits, such as rays or the typically SAR-dark parabolae seen on Venus, are observed.

This is a very small number of craters, far fewer than on airless, geologically inactive bodies. Titan's dense atmosphere should break up or slow small, incoming projectiles that would form craters of up to 20 m in diameter¹⁰. The presence of only one 450-km impact basin so far is not surprising; the distribution of such large craters on Rhea and Iapetus might imply several more comparably sized basins on Titan, though at this end of the size spectrum we expect very few¹¹. It is surprising, however, that we have not detected significant numbers of craters with diameters between 20 and 100 km. If Titan's cratering rate is similar to Rhea's, more than 200 craters larger than 20 km in diameter should have formed in the $\sim 2\%$ of the surface sampled using SAR. Their absence suggests that they have been erased or buried by tectonic and/or cryovolcanic processes. If Titan's crust is or was very thin, it is also possible that viscous relaxation made most such craters extremely shallow, and these were subsequently buried, perhaps by organic aerosols sedimenting from the stratosphere. All these possibilities require additional radar observations to cover more of the surface (and possibly find partially buried structures).

Fluvial features

A variety of apparent drainage patterns are observable in the T_3 swath, some lighter and some darker than their surround. They include linear channel-like features (hereafter "channels") 1–2 km wide and tens to hundreds of kilometers long, and SAR-bright formations, apparently dendritic drainages (hereafter "drainages"), at various scales close to Menrva. In contrast to the external channels, which are all SAR-bright, the short channels on the inner rim of the basin are both darker and brighter in SAR than the surrounding material. A well-developed 2- to 4-km-wide example of the latter appears to flow generally eastward and terminates at the southwestern rim. Others, less branched, seem to flow generally northeast

from just outside the southeast rim of the crater for about 200 km and terminate in a series of fan-shaped, delta-like forms that coalesce into a bright region, possibly a collecting basin. These channels include meandering, linear, and braided segments, with widths ranging from the limit of SAR resolution to 7 km. The arrangement of all these features suggests an E-to-NE downward regional slope. Several very different drainages appear along the inner rim of the crater, most prominently on the eastern side. These systems are short, 20–50 km long, tightly branched, with branches 1–2 km wide, extending roughly radially into the crater.

SAR backscatter of the large drainage patterns outside Menrva at various incidence angles is consistently 2 to 3 dB greater than that of the surrounding terrain. The SAR-bright area formed by coalescing channels east of the crater is large enough to be resolved by the Cassini imaging system and also appears bright at $0.93\ \mu\text{m}$ (ref. 6). The dark/bright modulation of the shorter channels strongly suggests resolved topographic slopes. Radarclinometric modelling and a consideration of the geometric distortion of these features indicate that they are 200–300 m deep, with inward-facing slopes of the order of 10° ; uncertainty in the assumed backscatter curve and the limited resolution of the images lead to errors of at most 50% in these depth estimates.

We suggest that the SAR-observed channels are produced by low-viscosity fluids, possibly liquid methane, as appears to be the case for smaller channels observed in the Huygens descent images¹². However, we cannot rule out participation of ammonia-water cryo-lava in the creation of the larger channels at temperatures enough above the eutectic to keep viscosity relatively low¹³. We regard ammonia-water fluid as a less likely agent than liquid hydrocarbons because we do not see evidence that the T_3 channels terminate in formations like those interpreted as flow lobes in the T_A data³. In the T_A channels, heating associated with volcanism might have created fractures along which liquid hydrocarbons flowed and carved channels, while in the T_3 region the causative agent of crustal fracturing was the large impact. Whether these processes also acted to liberate liquid hydrocarbons stored in the crust, or whether liquid was supplied by precipitation, cannot be determined from the data at hand.

Surficial dark streaks

Dark lineated streaks occur in regions of variable extent throughout the T_3 swath (Fig. 2a, f and g). The area of such regions ranges from $500\ \text{km}^2$ near the southeastern rim of the large basin to $35,000\ \text{km}^2$ in the western part of the swath, in total occupying approximately 10% of the T_3 swath. They appear in places to have an abrupt termination, possibly controlled by change of topography (for example, west of the basin), while in other places they blend and disappear in the surrounding surfaces (for example, the centre of the swath and west of the basin) or divert and wrap around the bright material of Sinlap. Individual streaks are ≤ 50 km in length, 500 m to 1 km across, and separated by 1–2 km. They are straight or sinuous, with prevalent orientation within 30° of east–west. Streak spacing for an informal sampling is 1.5 ± 0.4 km. The SAR expression of these features is typically as uniform dark bands, except in a few cases where there is some suggestion of a flanking bright edge.

The morphological characteristics and relationships of the dark streaks with the surrounding surfaces and orientation of Y-shaped junctions suggest they could represent linear (longitudinal) dunes formed by eastward aeolian (wind-generated) transport and accumulation, although they could as well be wind streaks with little or no topography. Longitudinal dunes imply a variable but typically bidirectional wind regime that is on average parallel to the dunes' long direction. SAR images of terrestrial dunes show similar patterns¹⁴. Remarkably similar images of snow dunes in Antarctica show alternating light and dark bands with narrow, low-backscatter bands alternating to wider, high-backscatter ones, which imply variation in grain size; these are also reported to have little topographic expression¹⁵. Radiometry observations of the Titan streaks are medium-bright, consistent with low-volume scattering, therefore suggesting a shallow-depth, grain-scale scattering effect.

Sources of sediments or particulate materials feeding aeolian dunes on Titan could include water ice, water-ammonia ice, solid hydrocarbons and nitriles such as acetylene, hydrogen cyanide, and other organic solid clathrates. It has been hypothesized¹⁶ that sand-sized material might be produced from a mix of such materials as

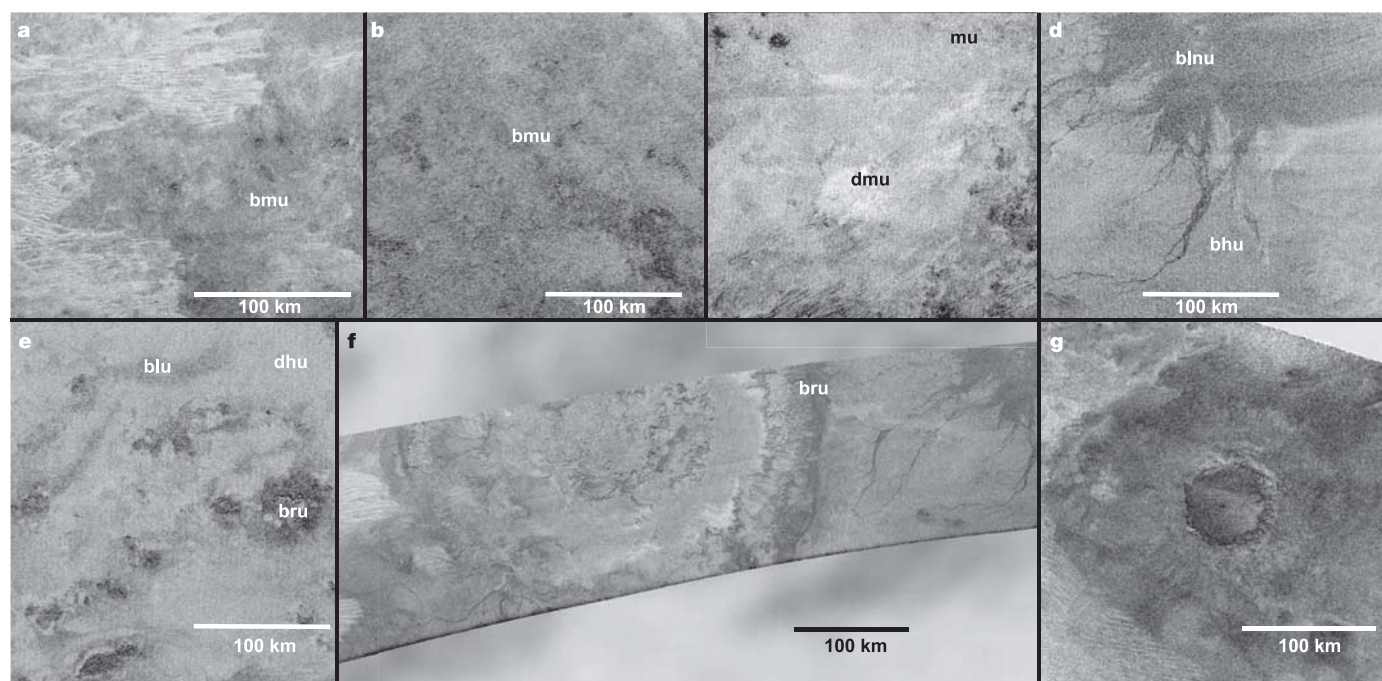


Figure 2 | Discrete mapping units in the T_3 SAR swath. Included are three units not detected in T_A : bmu (bright mottled unit), bhu (bright homogeneous unit) and bru (bright rough unit) (segments individually enhanced). **a** and **b**, bmu. **c**, mu (mottled unit); dmu (dark mottled unit).

d, bhu; **blnu** (bright lineated unit). **e**, bru; **blu** (bright lobate unit); **dhu** (dark homogeneous unit). **f**, 450-km impact basin Menrva (note ISS image in background). **g**, 80-km crater Sinlap.

impact ejecta; Huygens probe observations of channel networks suggest that fluvial erosion could also represent a viable source of sediments. Probe data indicate winds of $\sim 0.3 \text{ m s}^{-1}$ in the lowest part of its descent¹²; such wind speeds are comparable with the threshold speeds needed to transport (saltate) sand-sized particles in Titan's low gravity and thick atmosphere¹⁶. Winds of this magnitude have been predicted to be generated by Saturn's gravitational tide on Titan¹⁷ or by Titan's super-rotation, the resultant wind being almost entirely prograde (eastward). Cassini Imaging Subsystem (ISS) data have shown evidence of predominantly east–west aeolian features that suggest an eastward wind direction⁶.

Although dune patterns appear over a large portion of the T_3 swath, they do not appear in areas dissected by bright channels. The mobile dark material could also form the more amorphous dark patches observed in the radar data (for example, the dark mottled unit in Fig. 2).

Physical and electrical surface properties

To the SAR data, scatterometry and radiometry add information on the physical and electrical properties of the surface. T_3 scatterometry also observed the same terrain seen at the end of the T_3 SAR strip, yielding higher-resolution backscatter measurements. Relatively high-resolution radiometry data were obtained in both SAR and scatterometer passes between the active radar echoes.

A comparison of the common area covered by the scatterometer in T_A outbound and T_3 passes is presented in Fig. 3. Although there are several notable differences, the main one is that Sinlap is seen in the T_A outbound scan. Various specular scattering models have been successfully applied to planetary surfaces; although statistically acceptable fits result (see Fig. 3), the uncertainty in least-squares parameter estimates is large. We are not satisfied that the modelling of the current data establishes meaningful bounds on key physical quantities (for example, dielectric constant and root-mean-square slope). Nonetheless, it is clear that we are seeing a smooth surface component with low-to-moderate effective dielectric constants. We note that the scattering functions pictured in our plots are built up from measurements of diverse locations, in contrast with the

more-usual scatterometric observation of a specific location at multiple angles. Thus, in modelling, we are implicitly trying to average over potentially very diverse locations. Moreover, the nature of radar scattering from Titan may be intrinsically different from that encountered on other planetary radar targets.

Figure 1b shows the radiometric measurements obtained during the T_3 scatterometry pass and the T_A outbound pass. Regions of low radiometric brightness are strongly correlated with regions of high radar reflectivity in Fig. 3. Because the physical temperature variations are likely to be relatively small over this limited area, brightness variations can be ascribed to variations in emissivity rather than temperature. These in turn reflect variations in the physical or chemical properties of the surface. The observed radiometric temperature could result either from variations in the dielectric constant of near-surface material, or from variations in the degree of subsurface (volume) scattering. In the case of surface scattering, a high dielectric constant generally leads to high reflectivity and consequently low emissivity and brightness temperature. On the other hand, a lower surface dielectric constant can result in a decrease in brightness temperature if there is enhanced scattering from the subsurface. The correlation of radar reflection and low emissivity is most simply explained by the presence of enhanced volume scattering in the radar-reflective, radiometrically cool regions. However, this explanation is not unique, because the radiometrically dark regions could also have a higher dielectric constant to account for the lower emissivity and higher roughness to account for the radar return.

A feature associated with Sinlap is seen in both radiometry and scatterometry data and coincides approximately (within the radiometer/scatterometer resolution) with the bright halo. Figure 1c compares the eastern end of the T_3 SAR strip (left) with the accompanying radiometry (right). As discussed above, the decrease in emissivity of the ejecta could be explained by an increase in dielectric constant and the lack of volume scatter, suggesting, for example, an increase in dielectric constant from 2 to 3. Alternatively, the decrease in emissivity as well as the increased scattering seen both in the SAR swath and in the scatterometry could be due to increased

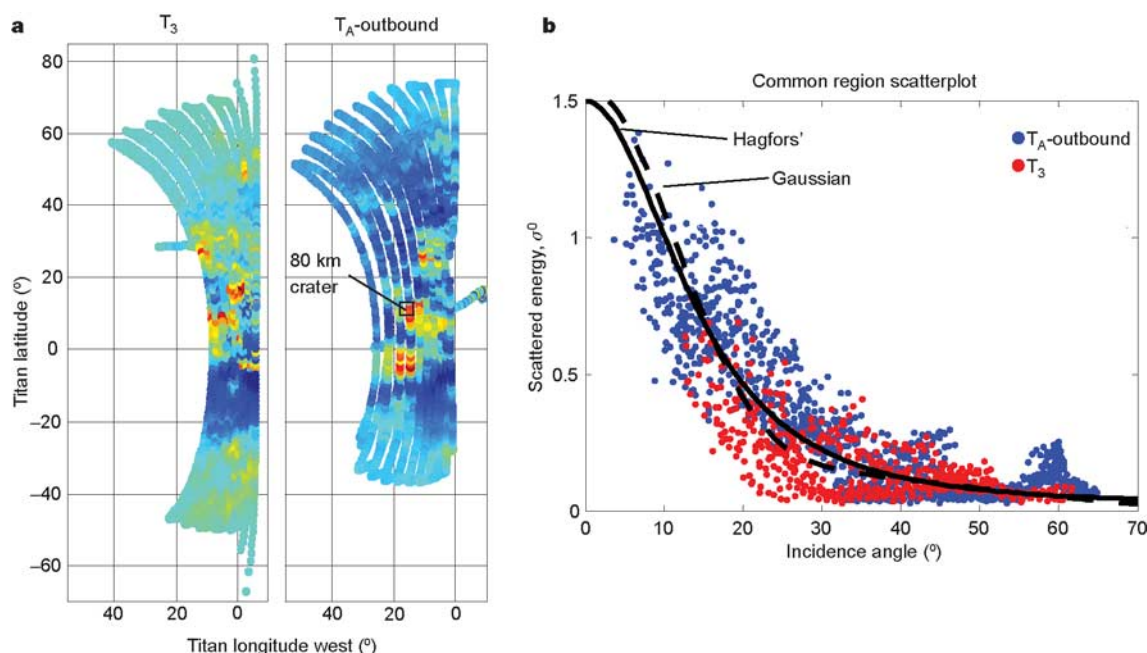


Figure 3 | Comparison of T_A outbound and T_3 scatterometry, showing consistency of measurements. **a**, The 80-km crater Sinlap is indicated. The plot of relative scattered energy σ^0 versus incidence angle from the common region in **b** is equally well fitted by several models. Pictured here are a

Hagfors specular scattering model, and a combination of a gaussian model and a cosine-like volume-scattering term. The Hagfors curve suggests a surface with a high dielectric constant (5), while the gaussian model implies a much lower value (<2).

volume scattering alone. More detailed modelling is needed to examine whether a similar dielectric constant change could explain the observed increase in SAR brightness as well, or whether volume scattering increases are required.

Discussion

The T_A and T_3 swaths cover approximately 3% of the surface of Titan. Two impact craters appear in the T_3 swath, whereas no unambiguous impact craters could be identified in the T_A swath. From the scarcity of 20–100 km craters observed, we conclude that the surface of Titan is predominantly young. The ubiquitous dark streaks in the T_3 swath we interpret as resulting from aeolian processes. Their presence indicates a relatively large amount of material capable of being transported and redeposited. This mobile material may be produced by fluvial or aeolian erosion or atmospheric deposition. As in the T_A swath, we see no clearly identifiable large bodies of liquid.

Both swaths provide ample evidence that Titan has been resurfaced recently in its geologic history. Both impact craters in the T_3 swath have been modified by erosion and infilling. The bright rough unit also provides support for modification and burial of older surfaces. Probable resurfacing processes include cryovolcanism, fluid flow, aeolian processes, and atmospheric deposition. In contrast to T_A , the T_3 swath contains no convincing examples of bright lobate flow features indicative of cryovolcanism. There are no structures similar to Ganesa Macula found in the T_A swath³. The T_3 swath is farther south than T_A , but at present there is insufficient coverage to determine whether cryovolcanic features are widespread and randomly distributed, or exhibit some geographic pattern.

Evidence for fluid flow appears in both swaths in sinuous, channel-like features, or fan-shaped deposits. A fluvial origin for these features has been bolstered by the identification of apparent drainage channels in the Huygens image data¹², although channels in the latter are at a much smaller scale than those in the T_A and T_3 swaths. We interpret the degradation of the southwest part of the impact basin rim as being caused by fluid flow.

Radar data of both T_A and T_3 swaths suggest that volume scattering may be occurring. Therefore, some of the features and textures in the swaths may be produced by subsurface changes in composition or roughness. The striking difference in overall appearance of the two swaths—with T_A seemingly dominated by endogenic processes, and T_3 more reflective of atmospheric and impact-related effects—emphasizes the diverse nature of Titan's surface and provides justification for the expectation of entirely new surface landforms yet to be detected in additional SAR images and other radar data.

Received 7 October 2005; accepted 4 April 2006.

1. Elachi, C. *et al.* Cassini Titan Radar Mapper. *Proc. IEEE* **79**, 867–880 (1991).
2. Elachi, C. *et al.* RADAR: The Cassini Titan Radar Mapper. *Space Sci. Rev.* **117**, 71–110 (2005).
3. Elachi, C. *et al.* Cassini radar views the surface of Titan. *Science* **308**, 970–974 (2005).
4. Sotin, C. *et al.* Release of volatiles from a possible cryovolcano from near-infrared imaging of Titan. *Nature* **435**, 786–789 (2005).
5. Mitri, G., Lunine, J., Showman, A. & Lopes, R. Resurfacing of Titan by ammonia-water cryomagma. *Icarus* (submitted).
6. Porco, C. *et al.* Imaging of Titan from the Cassini spacecraft. *Nature* **434**, 159–168 (2005).
7. Schenk, P. Central pit and dome craters: Exposing the interiors of Ganymede and Callisto. *J. Geophys. Res.* **98**, 7475–7498 (1993).
8. Kirk, R. *et al.* & the Cassini Radar Team. RADAR reveals Titan topography. *Proc. 36th Annu. Lunar Planet. Sci.* XXXVI abstr. 2227 (2005).
9. Melosh, H. *Impact Cratering: A Geologic Process* 87–111 (Oxford Univ., Oxford, 1989).
10. Lorenz, R. Impacts and cratering on Titan: A pre-Cassini view. *Planet. Space Sci.* **45**, 1009–1019 (1997).
11. Wood, C., Lopes, R., Stofan, E., Paganelli, F. & Elachi, C. Impact craters on Titan? Cassini RADAR view. *Proc. 36th Annu. Lunar Planet. Sci.* XXXVI abstr. 1117 (2005).
12. Tomasko, M. G. *et al.* Rain, winds and haze during the Huygens probe's descent to Titan's surface. *Nature* **438**, 765–778 (2005).
13. Kargel, J., Croft, S., Lunine, J. & Lewis, J. Rheological properties of ammonia-water liquids and crystal-liquid slurries: Planetological applications. *Icarus* **89**, 93–112 (1991).
14. Blom, R. & Elachi, C. Spaceborne and airborne imaging radar observations of sand dunes. *J. Geophys. Res.* **86** (B4), 3061–3073 (1981).
15. Fahnestock, M., Scambos, T., Shuman, C., Arthern, R., Winebrenner, D. & Kwok, R. Snow megadunes fields on the East Antarctica Plateau: Extreme atmosphere-ice interaction. *Geophys. Res. Lett.* **27**(22), 3719–3722 (2000).
16. Lorenz, R., Lunine, J., Grier, J. & Fisher, M. Prediction of aeolian features on planets: Application to Titan paleoclimatology. *J. Geophys. Res.* **100**, 26377–26386 (1995).
17. Tokano, T. & Neubauer, F. Tidal winds on Titan caused by Saturn. *Icarus* **158**, 499–515 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge those who designed, developed, and operate the Cassini/Huygens mission. Cassini is a joint endeavour of the National Aeronautics and Space Administration (NASA), the European Space Agency (ESA), and the Italian Space Agency (ASI) and is managed by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

Author Contributions C.E. is the RADAR Team Leader. All authors contributed equally to this work.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.W. (steve.wall@jpl.nasa.gov).

ARTICLES

Stromatolite reef from the Early Archaean era of Australia

Abigail C. Allwood^{1,2}, Malcolm R. Walter^{1,2}, Balz S. Kamber³, Craig P. Marshall^{1,4} & Ian W. Burch²

The 3,430-million-year-old Strelley Pool Chert (SPC) (Pilbara Craton, Australia) is a sedimentary rock formation containing laminated structures of probable biological origin (stromatolites). Determining the biogenicity of such ancient fossils is the subject of ongoing debate. However, many obstacles to interpretation of the fossils are overcome in the SPC because of the broad extent, excellent preservation and morphological variety of its stromatolitic outcrops—which provide comprehensive palaeontological information on a scale exceeding other rocks of such age. Here we present a multi-kilometre-scale palaeontological and palaeoenvironmental study of the SPC, in which we identify seven stromatolite morphotypes—many previously undiscovered—in different parts of a peritidal carbonate platform. We undertake the first morphotype-specific analysis of the structures within their palaeoenvironment and refute contemporary abiogenic hypotheses for their formation. Finally, we argue that the diversity, complexity and environmental associations of the stromatolites describe patterns that—in similar settings throughout Earth’s history—reflect the presence of organisms.

The oldest identifiable fossil assemblages from the Earth’s early biosphere are crucial to our understanding of the origins of life on our planet and, by analogy, how and where we might find traces of extinct primitive life on other planets and moons. The SPC’s widespread, well-preserved and morphologically diverse stromatolitic outcrops offer unique, ecosystem-scale insights to such questions. The formation was described almost three decades ago as a partly evaporitic, shallow-water succession containing conical structures of possible microbial origin (stromatolites)^{1,2}. However, subsequent research indicated that abiogenesis could not be excluded³. More convincing evidence of microbial sedimentation was later found in a small outcrop of well-preserved conical stromatolites at the ‘Trendall Locality’⁴. Complex morphological attributes were cited as compelling evidence for biogenesis. However, a subsequent study at another location argued that the complexity of the Trendall stromatolites was rare in the formation⁵. Those authors interpreted a hydrothermal environment of deposition for the SPC and described the ‘stromatolite’ laminae as isopachous, suggesting that the structures were abiotic hydrothermal precipitates. Another study cited detrital sedimentary structures and seawater rare-earth-element (REE) chemistry of carbonate in the Trendall area as evidence that stromatolite accretion occurred by microbial trapping and binding of grains in a marine environment⁶. However, mathematical models show that abiotic marine cementation of detritus can also produce stromatolite-like structures^{7,8}. Thus, the origin of the SPC stromatolites remains ambiguous because of contrasting observations of stromatolite morphology drawn from disparate localities and diverging environmental interpretations.

To address this conundrum we investigated stromatolite morphologies against the backdrop of their palaeoenvironmental setting, across more than 10 km of relatively continuous outcrop where the formation is best preserved (Supplementary Fig. S1). For simplicity we adopt the term ‘stromatolite’ (a term implying biogenesis⁹) but acknowledge the as yet unproven biogenicity. Here we show that

seven separate stromatolite facies arose in different environments and developmental stages of an isolated peritidal carbonate platform during transgression of previously emergent crust. We argue that the SPC is most plausibly interpreted as a fossil microbial carbonate platform on the following grounds: first, the palaeoenvironmental setting; second, combined attributes of the stromatolites; third, stromatolite distribution within the palaeoenvironment; and last, similarity with younger microbially mediated peritidal carbonates.

Stratigraphy and facies

The SPC is a sedimentary rock succession deposited unconformably over at least 3.43-Gyr-old volcanic/sedimentary rocks and overlain by 3.350–3.315-Gyr-old volcanic rocks^{10,11}. The succession is subdivided in the study area into four members (Supplementary Fig. S2). The basal member (M1) is a thin, discontinuous jasper/chert boulder conglomerate layer. Member 2 (M2) is a layer of laminated stromatolitic carbonate/chert about 10–20 m thick, subdivided into three beds. Member 3 (M3) consists of bedded black chert, and member 4 (M4) is a silicified fining-upward clastic/volcaniclastic unit. To the south, M2–M4 merge into a succession of laminated chert more than 70 m thick with a thin basal layer of wavy laminated chert.

We identify six different stromatolite types in M2, consisting of pseudocolumnar structures composed of, and linked by, rhythmic chert/dolomite laminae. A seventh stromatolite facies occurs in M3 and is composed of laminated ironstone in chert. The SPC records evolving environments that controlled stromatolite development laterally and through time, from an initial transgressive rocky coast setting (M1); to rapid development of an isolated, southward-deepening peritidal stromatolitic carbonate platform during continued transgression (M2); and finally, resumption of hydrothermal activity, volcanism and clastic sedimentation in a southward-subsiding basin (M3–M4).

The interpretation of M1 as a rocky shoreline deposit hinges on the following: first, the wide but laterally discontinuous distribution of

¹Australian Centre for Astrobiology/Macquarie University Biotechnology Research Institute, and ²Department of Earth and Planetary Sciences, Macquarie University, Herring Road, Sydney, New South Wales 2109, Australia. ³Department of Earth Sciences, Laurentian University, 933 Ramsey Lake Road, Sudbury, Ontario P3E 6B5, Canada. ⁴Vibrational Spectroscopy Facility, School of Chemistry, The University of Sydney, Sydney, New South Wales 2006, Australia.

boulders; second, rounded, clast-supported fabric, indicating energetic deposition; third, correlation between clast and substrate lithologies, indicating minimal clast transport; fourth, stratigraphic position over an unconformity; fifth, substrate-dependent lateral transition from clustered and isolated large boulders on headland-type 'wave-cut' platforms to embayment beach-type conglomerates; sixth, the presence of palaeo-cliffs, fissures and cavities in the substrate (namely seacliffs and other shoreline erosion features); and last, soft mud intraclasts and desiccation cracks associated with local mudstone substrate, similar to weathering features in Holocene mudstones exposed in modern intertidal zones¹² (Supplementary Information).

At the M1 upper contact, laminated carbonates of M2 encrust the M1 boulders, 'cementing' many in precarious positions. Carbonate intraclasts occur locally within the upper part of M1, and a single layer of isolated chert boulders occurs locally about 80 cm above the contact (Supplementary Figs S4–S6). Thus, the contact is transitional, demonstrating depositional continuity from M1 to M2 and implying a shallow-water depositional setting for M2. Significantly, the contact marks a sudden change from very coarse terrigenous clastic sediment to finely laminated dolomite-chert virtually free of terrigenous sediment. The rapid switch is difficult to account for in a non-marine setting and is most plausibly explained by sudden drowning of an isolated low-relief landmass by a shallow sea. A marine environment for M2 is confirmed by the primary seawater REE patterns of the dolomite and chert (Supplementary Information). More specifically, a peritidal marine carbonate platform environment is indicated by a suite of interrelated observations (described below) on the facies associations and architecture of M2.

As is characteristic of peritidal carbonates¹³, M2 records a series of minor regressions within an overarching trend of rising sea level. The three regressive cycles are recorded as three stacked beds (Supplementary Figs S2 and S6) capped by locally eroded layers of evaporite crystal pseudomorphs. Each bed has a deeper-water facies association—and different stromatolites—in comparison with the bed below.

Bed 1 of M2 (first regressive cycle) is characterized by rapid lateral facies and thickness variations controlled by underlying topography. Where bed 1 overlies conglomerate on chert substrate, laminated carbonate encrusts the boulders and forms striking, upward-expanding, domical-laminated pseudocolumns. Where M1 conglomerate is absent, non-encrusting domical pseudocolumns occur. These 'encrusting and domical laminites' (stromatolite facies 1; Fig. 1) pass laterally to flat or irregular laminated and fenestral carbonate, notably where the underlying rocks are bedded conglomerate over mudstone (that is, a probable drowned embayment/low area). The amplitude of individual encrusting/domical laminae indicates minimum water depths of about 1 m in the lower part of bed 1. The upper layers of bed 1 mark a gradual infilling of topography and transition to centimetre-scale 'small crested/conical laminites' (stromatolite facies 2; Fig. 1), or flat laminite, with scattered evaporite crystal casts and intraclast conglomerates (storm and tidal channel deposits)—a characteristic intertidal to supratidal carbonate facies association¹³. Bed 1 is capped throughout the area by a 0.5–1.8-m-thick layer of evaporite crystal pseudomorphs, indicating regression and widespread restricted circulation. Local erosion surfaces, solution cavities and collapse breccia (karst) at the upper contact, particularly in the northern area (Supplementary Figs S6 and S13), coupled with the trend of decreasing bedform amplitude and increasing evaporite deposition, indicate shoaling conditions through the deposition of bed 1.

Bed 2 of M2 displays less pronounced, broader-scale facies variations than bed 1, reflecting continued carbonate build-up, transgression and topographic infill. The facies association consists of four different types of stromatolites—'egg-carton laminites' and 'large complex cones' (LCC) in the central to southern area; and in the north, large 'cusped swales' grading up to 'small crested/conical

laminites' (Fig. 1)—intercalated with flat laminite and edgewise intraclast conglomerates (storm/tidal channel deposits). The flat laminite displays low-angle downlap, cross-lamination, scours, intraclast microconglomerates and graded lamination, indicating the deposition of granular sediment. Rare desiccation cracks and structures resembling blister mats or desiccated microbial mats in the northern area indicate intermittent exposure (Supplementary Information). That facies association is typical of intermittently exposed, intertidal to lower supratidal carbonates¹³. As with bed 1, shoaling is reflected in an up-section trend of decreasing bedform amplitudes, increasing abundance of evaporites and a widespread, locally eroded capping evaporite layer. However, erosion features are less prominent than in bed 1, indicating that regression was less pronounced at the end of the second regressive cycle.

Bed 3 of M2 is a relatively uniform, extensive layer of 'wavy laminites'—a typical subtidal carbonate facies¹³—with abundant laminar onlap indicating possible loose sediment deposition in water continuously agitated by waves and currents. However, the structures are commonly cone-shaped with slopes up to about 50° above horizontal, indicating possible accretionary mediation and prompting their classification as stromatolite facies 6 (Supplementary Information). Local breccia deposits containing boulders of the underlying carbonate facies indicate erosion at the end of the third regressive cycle.

M3, which sharply overlies M2, is a layer of bedded carbonaceous chert about 1–3 m thick with thin white quartz bands, silicified pebble conglomerates and rare silicified evaporite beds. Thin, continuous, bumpy/wavy layers of laminated ironstone (stromatolite facies 7; Fig. 1) occur at irregular stratigraphic intervals through the chert (Supplementary Information). These observations indicate possible shallow, silica-depositing and clastic sediment-depositing, partially restricted conditions.

M4 is a clastic sediment succession of variable thickness, deposited in subsiding fault blocks and fining upwards from conglomerate to bedded, silicified tuffaceous mudstone. The presence of black chert veins terminating in zones of phreatomagmatic brecciation in the tuffaceous mudstone, as well as hydrothermal REE signatures of the breccia matrix, vein and bedded cherts, indicate the onset of hydrothermal activity during the deposition of M4 (Supplementary Information).

In the far south, the SPC thickens to more than 70 m and consists almost entirely of flat laminated black/grey chert with carbonate silt (Supplementary Figs S2 and S14). Only a thin layer (less than 2 m thick) of wavy laminite occurs near the base, locally overlying chert conglomerate (M1). Convolute laminae (namely slump folds; Supplementary Fig. S15) indicate deposition on a slope dipping approximately south. The laminites are interpreted as platform slope deposits: the basal wavy laminite and the lower laminite succession probably correspond to M2, as indicated by the abundant carbonate silt; the remaining section is interpreted as the equivalent of M3–M4.

On the basis of observations in the four members, we propose that the SPC stromatolites formed on an isolated, transgressive peritidal carbonate platform that developed on a drowned landmass as soon as erosion surfaces and rocky coastal deposits became submerged and clastic influx ceased. Evidence for periodic restriction and exposure increases northward, reflecting the on-platform direction. Southward facies trends confirm basinward deepening to the south. Abundant intraclast conglomerates, scours and ripped-up laminites in M2 in the Trendall area indicates that a platform margin may have absorbed high wave energy there at times. Onset of stromatolite development coincided with the end of siliciclastic sediment influx, and the end of stromatolite formation coincided with the resumption of siliciclastic sedimentation, hydrothermal activity and volcanism. The stromatolites were spatially restricted to shallow-water areas.

Origin of stromatolites

This study introduces new evidence relevant to the debate about the

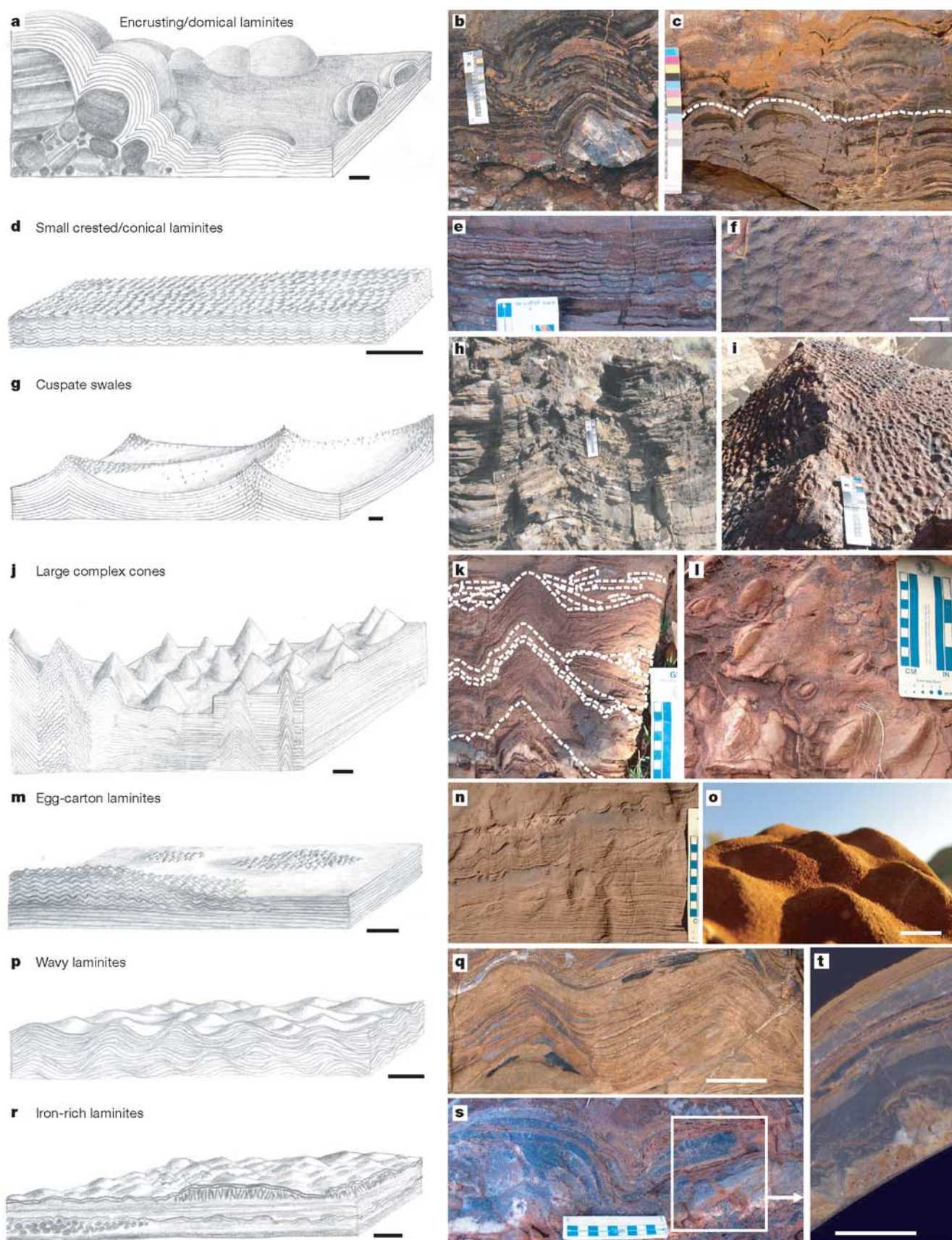


Figure 1 | Stromatolite facies of the Strelley Pool Chert, showing reconstructed three-dimensional views and outcrop photographs. a–c, ‘Encrusting/domical laminites’; d–f, ‘small crested/conical laminites’; g–i, ‘cuspate swales’; j–l, ‘large complex cones’ (dashed lines in k trace lamina shape and show outlines of intraclast conglomerate piled against the cone at

two levels). m–o, ‘Egg-carton laminites’; p, q, ‘wavy laminites’; r–t, ‘iron-rich laminites’ (t is a cut slab). The scale card in b, h and i is 18 cm. The scale card increments in c, e, k, l, n and s are 1 cm. The scale bar in o is about 1 cm. The scale bars in the remaining pictures are about 5 cm.

biogenicity of SPC stromatolites. Not one but seven stromatolite types occur, each with distinct morphological attributes (Fig. 1, Supplementary Fig. S16 and Supplementary Tables S1–S8) and distribution. All have syndepositional origins and possess geometries that are inconsistent with purely mechanical deposition (Supplementary Information). The palaeoenvironmental context does not support a general interpretation of the stromatolites as abiogenic hydrothermal precipitates⁵. Indeed, the identification of multiple stromatolite facies with distinct attributes means that no generalized abiogenic interpretation can be applied without differentiating between attributes of individual facies. Finally, the wealth of information contained in the seven facies introduces a degree of complexity not previously considered in growth models for such stromatolites^{7,14,15}. These factors provide significant impetus for a re-examination of the SPC stromatolites by facies type and as facies associations within the palaeoenvironmental framework.

We first examine the LCC stromatolites—which include previously described ‘Trendall Locality’ structures⁴—and consider whether they are more plausibly explained as biogenic or abiogenic (chemically precipitated) accretionary structures. In a purely chemical sedimentary regime it is first difficult to explain why laminae took on a specific and strongly inherited conical shape at certain heterogeneously spaced points, whereas everywhere else they remained relatively flat. Uneven spacing can occur in abiogenic accretionary structures, according to theoretical models^{7,8}, but not in combination with the particular conical morphology of the LCC structures⁸. Purported abiogenic structures with conical laminae have been described from the geological record—but they differ in morphology from the LCCs owing to the surface-normal growth and isopachous geometry of the crystalline layers, which produce pseudocolumns with upward-diverging margins¹⁶ (Supplementary Fig. S19). Previous studies proposed that SPC stromatolite laminae are isopachous and are therefore chemically precipitated⁵. However, our measurements show that the laminae are not isopachous. They thin on the cone slopes by the exact amount that ensures that the vertical ‘depth’ of the laminae is consistent throughout, the cone dimensions are inherited through successive layers, and a palaeovertical, parallel-sided pseudocolumn is formed (Supplementary Fig. S18). Isopachous or any laminar geometry other than that measured will alter the cone size and/or shape through successive layers (Supplementary Fig. S19). Thus, the palaeovertical conical pseudocolumn is a structure requiring specific, vertically directed accretion^{8,14}; it has no known abiogenic analogue or model⁸. Palaeovertical growth is echoed in the vertical alignment of LCC pseudocolumns on palaeoslopes¹⁰ (Supplementary Fig. S20) and is most plausibly explained—and demonstrated in modern conical stromatolites¹⁷—as the product of upward-migrating microbial colonies at the sediment/water interface.

In addition to geometric arguments for biogenesis, textural differences between cone and interspace laminae indicate that unique, localized sedimentary processes may have prevailed on the cones. In contrast with cone laminae, interspace laminae have graded fabrics and onlapping or infilling geometries. They also have more variable thickness and measurably different fabric—on average 45% are tangentially truncated, in comparison with less than 5% on the cones (Supplementary Information). That indicates a dominantly mechanical deposition of grains in the interspaces and a different combination of processes and/or sediment types on the cones, most plausibly explained by microbial influence (for example, trapping and binding and/or intra-mat mineral precipitation) at the cone surfaces.

Finally, a significant trace chemical signature detected in the LCC structures that is only known to occur in biosediments is the 250-fold enrichment of REE in carbonate laminae relative to chert laminae (Supplementary Fig. S22). Previous geochemical studies have indicated that microbial carbonates have REE concentrations consistently 200–300-fold higher than their syndepositional and early diagenetic abiotic counterparts^{18–20}.

Thus, the LCC structures have geometric, textural and chemical attributes that provide compelling evidence for biogenesis, and have no natural or experimental abiogenic analogue. The LCC structures are most plausibly explained as biosedimentary structures that resemble younger conical stromatolites of widely accepted biogenicity, such as *Conophyton*²¹ and related forms.

The other six stromatolite facies also display attributes that indicate probable biogenesis (Supplementary Information). The encrusting/domical laminates have granular sediment in their laminae (Supplementary Fig. S17), indicating the likely presence of a microbial mat to trap and bind grains—or to form grains *in situ*—and precluding their interpretation as purely chemical precipitates. Accretion of the sediment into domical laminae is more plausibly interpreted as a biogenic feature, shared by known stromatolite taxa such as *Colonnella*²² and *Conusella*²³. The wavy laminates look similar to climbing ripples in cross section; however, the conical lamina shapes and steep slopes indicate microbial influence and are comparable with the stromatolite taxon *Irregularia*²⁴. Egg-carton laminates, like the LCC structures, are most plausibly interpreted as having been formed by biogenic processes similar to those that formed conical stromatolites such as *Conophyton*. The cusped swales are perhaps least compatible with the abiogenic hypothesis: their laminae form strongly inherited, highly complex metre-scale networks of crested ridges with concave-sided pseudo-cones at the ridge intersections and smaller cones commonly adorning the crests. To have been formed abiotically, their morphology would have required an improbable combination of physical and chemical processes, and even more improbable conditions to maintain those processes through the deposition of scores of near-identical laminae. In contrast, the biogenic hypothesis is plausible—even highly probable—given that similar morphological attributes occur in known stromatolite taxa such as *Thesaurus*²⁵. Perhaps the most significant observation relates to slumped cones on the flanks of some cusped swales: these indicate that the sediment was coherent but soft, and thus that coherence was probably provided by flexible microbial mats rather than rigid crystalline crusts (Supplementary Information). Abiogenic replication of the complex morphology, the inferred flexible but cohesive laminar rheology, morphological inheritance through metres of finely laminated sediment, inhomogeneous adornment with small cones, crystals and bumps, and repetition of the morphology over a wide area is so complex as to be implausible.

Thus, seven distinct stromatolite facies in the SPC each display suites of attributes that are readily explained by biogenesis and are present in recognized stromatolite taxa from the geological record. In contrast, to interpret each facies as abiogenic requires not only a balance of processes that is unknown and unlikely in the natural world, but also lateral and temporal persistence of those processes to reproduce the vertically inherited, laterally repeated stromatolite morphologies. Perhaps the most compelling point is that—for the entire SPC carbonate platform to be interpreted as abiogenic—not one, but seven different unusually balanced sets of abiogenic processes must have operated discretely, persistently and sometimes simultaneously on the platform.

Finally, not only do the individual and collective stromatolite morphologies indicate biogenesis, but several factors are consistent with ecologically controlled growth of a microbial reef: especially the diversity of stromatolites—each with features resembling known microbialites and none resembling any known abiogenic structure, their occurrence in a transgressive carbonate platform deposit, their association with different platform environments, and the absence of stromatolites in deeper water.

This array of constraining factors strongly indicates that organisms flourished on a broad peritidal platform 3.43 Gyr ago in the Pilbara, rapidly taking hold and creating a reef-like build-up in shallow waters as surfaces became submerged. The variety of stromatolites present indicates that the SPC may contain not only some of Earth’s earliest fossils but also a diverse fossil ‘ecosystem’, sustained by

shallow seawater free of terrigenous influx—ideal conditions for phototrophism and a consequence of the location on a drowned topographic high. Moreover, the stratigraphic position of the SPC between thick igneous and hydrothermal successions shows that, although early life may have gained a foothold much earlier in ‘extreme’ volcanic²⁶ or hydrothermal²⁷ environments, the oldest surviving indications of a firm foothold are associated with a rare pause in igneous and hydrothermal activity, and with the onset of relatively ‘normal’ shallow marine conditions similar to those that have nurtured marine biodiversity throughout geological history. Perhaps in this instance, abiogenesis is the ‘extraordinary claim’ that requires extraordinary proof, whereas biogenesis offers an ordinary, plausible explanation for the nature of the SPC stromatolitic carbonate platform.

Received 9 September 2005; accepted 29 March 2006.

1. Lowe, D. R. Stromatolites 3,400 Myr old from the Archean of Western Australia. *Nature* **284**, 441–443 (1980).
2. Lowe, D. R. Restricted shallow water sedimentation of Early Archean stromatolitic and evaporitic strata of the Strelley Pool Chert, Pilbara Block, Western Australia. *Precamb. Res.* **19**, 239–283 (1983).
3. Lowe, D. R. Abiotic origin of described stromatolites older than 3.2 Ga. *Geology* **22**, 387–390 (1994).
4. Hofmann, H. J., Grey, K., Hickman, A. H. & Thorpe, R. I. Origin of 3.45 Ga coniform stromatolites in Warrawoona Group, Western Australia. *Geol. Soc. Am. Bull.* **111**, 1256–1262 (1999).
5. Lindsay, J. F. *et al.* The problem of deep carbon—an Archean paradox. *Precamb. Res.* **143**, 1–22 (2005).
6. Van Kranendonk, M. J., Webb, G. E. & Kamber, B. S. Geological and trace element evidence for a marine sedimentary environment of deposition and biogenicity of 3.45 Ga stromatolitic carbonates in the Pilbara Craton, and support for a reducing Archean ocean. *Geobiology* **1**, 91–108 (2003).
7. Grotzinger, J. P. & Rothman, D. H. An abiotic model for stromatolite morphogenesis. *Nature* **383**, 423–425 (1996).
8. Grotzinger, J. P. & Knoll, A. Stromatolites in Precambrian carbonates; evolutionary mileposts or environmental dipsticks? *Annu. Rev. Earth Planet. Sci.* **27**, 313–358 (1999).
9. Awramik, S. M., Margulis, L. & Barghoorn, E. S. in *Stromatolites* (ed. Walter, M. R.) 149–162 (Elsevier, Amsterdam, 1976).
10. Van Kranendonk, M. J., Hickman, A. H., Smithies, R. H. & Nelson, D. R. Origin and evolution of the Archean North Pilbara Terrain, Pilbara Craton, Western Australia. *Econ. Geol.* **97**, 605–732 (2002).
11. Van Kranendonk, M. J. & Pirajno, F. Geochemistry of metabasalts and hydrothermal alteration zones associated with c. 3.45 Ga chert and barite deposits; implications for the geological setting of the Warrawoona Group, Pilbara Craton, Australia. *Geochem. Exp. Environ. Anal.* **4**, 253–278 (2004).
12. Knight, J. Processes of soft-sediment clast formation in the intertidal zone. *Sedim. Geol.* **181**, 207–214 (2005).
13. Flugel, E. *Microfacies of Carbonate Rocks* (Springer, Berlin, 2004).
14. Bachelor, M. T., Burne, R. V., Henry, B. I. & Watt, S. D. Deterministic KPZ model for stromatolite laminae. *Physica A (Amsterdam)* **282**, 123–136 (2000).
15. Jogi, P. M. & Runnegar, B. Quantitative methods for evaluating the biogenicity of fossil stromatolites. *Astrobiology* **5**, 293 (2005).
16. Pope, M. & Grotzinger, J. P. in *Carbonate Sedimentation and Diagenesis in the Evolving Precambrian World* (eds Grotzinger, J. P. & James, N. P.) 103–122 (SEPM (Soc. Sedim. Geol.) Spec. Publ. 67, Tulsa, Oklahoma, 2000).
17. Walter, M. R., Bauld, J. & Brock, T. D. in *Stromatolites* (ed. Walter, M. R.) 273–310 (Elsevier Science, Amsterdam, 1976).
18. Webb, G. E. & Kamber, B. S. Rare earth elements in Holocene reefal microbialites: a new shallow seawater proxy. *Geochim. Cosmochim. Acta* **64**, 1557–1565 (2000).
19. Kamber, B. S. & Webb, G. E. The geochemistry of late Archean microbial carbonate: implications for ocean chemistry and continental erosion history. *Geochim. Cosmochim. Acta* **65**, 2509–2525 (2001).
20. Nothdurft, L. D., Webb, G. E. & Kamber, B. S. Rare earth element geochemistry of Late Devonian reefal carbonates, Canning basin, Western Australia: confirmation of a seawater REE proxy in ancient limestones. *Geochim. Cosmochim. Acta* **68**, 263–283 (2004).
21. Komar, V. P., Raaben, M. E. & Semikhatov, M. A. Conophytos in the Riphean of the USSR and their stratigraphic significance. [in Russian] *Trans. Geol. Inst. Akad. Nauk. SSSR* **131**, 1–73 (1965).
22. Komar, V. A. Columnar stromatolites from the north of the Siberian Platform. [in Russian] *Uch. Zap. Nauchno-Issled. Inst. Arkt. Paleontol. Biostratigr.* **6**, 84–105 (1964).
23. Golovanov, N. P. *Opornyy razrez verkhnedokembriyskikh otlozheniy zapadnoggo sklona of Anabarskogo podnyatiya (Sbornik Statey)* 60–89 (Nauchno-Issled. Inst. Geol. Arkt., Leningrad, 1970).
24. Korolyuk, I. K. Stromatolites of the Lower Cambrian and Proterozoic of the Irkutsk Ampitheatre. [in Russian] *Trans. Inst. Geol. Razrab Goryuchikh Iskop. Akad. Nauk SSSR (Moscow)* **1**, 112–167 (1960).
25. Vlasov, F. Ya. in *Materials on the Middle Palaeozoic Palaeontology of the Urals and Siberia* [in Russian] 101–124 (Ural. Sci. Centre Akad. Nauk SSSR, Sverdlovsk, 1977).
26. Banerjee, N. R., Furnes, H., Muelenbachs, K., Staudigel, H. & DeWit, M. Preservation of ~3.4–3.5 Ga microbial biomarkers in pillow lavas and hyaloclastites from the Barberton Greenstone Belt, South Africa. *Earth Planet. Sci. Lett.* **241**, 707–722 (2006).
27. Corliss, J. B. *et al.* An hypothesis concerning the relationship between submarine hot springs and the origin of life on Earth. *Oceanol. Acta* **4**, 59–69 (1981).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Van Kranendonk, S. Awramik and K. Grey for discussion and reviews; A. Grieg for assistance with REE analyses; the Geological Survey of Western Australia (GSWA) and the Pilbara Regiment for field support; and the NSW Dept of Commerce Heritage Restoration Services for large sample cutting. Research was supported by Macquarie University and the M.U. Biotechnology Research Institute. A.C.A. was supported by an Australian Postgraduate Award. All samples were collected with GSWA permission.

Author Contributions A.C.A. undertook the study and wrote the paper with supervision from M.R.W. and C.P.M. I.W.B. gave field assistance, and research and writing assistance. B.K. supervised and assisted with REE analyses.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.C.A. (aallwood@els.mq.edu.au).

Noise-resistant and synchronized oscillation of the segmentation clock

Kazuki Horikawa^{1*}, Kana Ishimatsu^{1*}, Eiichi Yoshimoto^{2*}, Shigeru Kondo² & Hiroyuki Takeda¹

Periodic somite segmentation in vertebrate embryos is controlled by the 'segmentation clock', which consists of numerous cellular oscillators. Although the properties of a single oscillator, driven by a *hairy* negative-feedback loop, have been investigated, the system-level properties of the segmentation clock remain largely unknown. To explore these characteristics, we have examined the response of a normally oscillating clock in zebrafish to experimental stimuli using *in vivo* mosaic experiments and mathematical simulation. We demonstrate that the segmentation clock behaves as a coupled oscillator, by showing that Notch-dependent intercellular communication, the activity of which is regulated by the internal *hairy* oscillator, couples neighbouring cells to facilitate synchronized oscillation. Furthermore, the oscillation phase of individual oscillators fluctuates due to developmental noise such as stochastic gene expression and active cell proliferation. The intercellular coupling was found to have a crucial role in minimizing the effects of this noise to maintain coherent oscillation.

Synchronized oscillation is a universal feature of many biological systems, such as the fluorescent 'blinking' of fireflies or the rhythmic cellular activities of the neuronal and cardiac systems^{1,2}. A transcriptional oscillator that is responsible for vertebrate somitogenesis, known as the segmentation clock, also displays synchronized oscillation (Fig. 1 and Supplementary Notes S1)^{3–5}. The segmentation clock is an ensemble of numerous cellular oscillators, located in the unsegmented pre-somitic mesoderm (PSM). The travelling wave generated from the synchronized oscillation in the posterior PSM sweeps anteriorly and arrests around the future segmentation point during the formation of each somite (every 30 min in zebrafish; see also Supplementary Notes S1)^{3,6}. It is believed that the frequency of this clock oscillation controls the size of the resulting segments (Supplementary Notes S2)^{7–9}, although no direct evidence for this has been shown. Previous genetic analyses suggested that the segmentation clock can be considered as a coupled oscillator, wherein numerous cellular oscillators, driven by negative-feedback regulation of *hairy*-related transcriptional repressors^{10–14}, are coupled by oscillator-linked Notch-mediated intercellular communication (Box 1 and Supplementary Notes S1)^{15,16}. However, because of the complex network of genes that function both intra- and intercellularly, we are still far from having a complete understanding of the mechanisms controlling the collective behaviour of the segmentation clock (that is, the synchronized oscillations, or the travelling and arresting waves, including their transition in the oscillation modes).

In the present study we have focused on the mechanisms of synchronized oscillation, which is the earliest and most basic mode of oscillation generated in the posterior PSM (Fig. 1). In particular, we asked how the intercellular communication coordinates individual cellular oscillators and maintains a global oscillation pattern. By using zebrafish embryos, we experimentally and mathematically analysed the system-level response of the normal oscillator to experimental stimuli.

Acceleration of clock oscillation induced by *her*-MO cells

To understand better the function of Notch signalling in the

segmentation clock, we examined how the normal oscillators respond to activated Notch signalling. In this assay, we unilaterally transplanted genetically modified cells at the blastula stage into the marginal mesoderm that was fated to become the somite, then analysed the effects of explants upon both the segment positions and the oscillation pattern of *her1* (zebrafish *hairy-related gene1*) at the segmentation stage (Fig. 2a). We first examined the effect of non-oscillating cells with high levels of Notch ligands on host segmentation. These donor cells were obtained from embryos in which the translation of *her1* and *her7* had been inhibited by morpholino anti-sense oligonucleotides (*her*-MO cells). *her*-MO cells are expected to continuously activate Notch signalling in surrounding cells, because the expression of *deltaC* is upregulated due to the absence of Her1/7-dependent repression (Supplementary Notes S1)^{10,14}. The results of this experiment were striking: in most of the transplanted embryos (91 out of 96) the segment positions were anteriorly shifted on the transplanted side (Fig. 2b). In addition, the segment size was locally reduced around the explants. This segment-shift activity of *her*-MO cells was also found to depend upon the function of DeltaC, as its depletion in donor cells abolishes the segment-shift activity of *her*-MO cells (Fig. 2c). This is also the case for DeltaD¹⁷, another ligand for Notch broadly expressed in the posterior PSM (data not shown).

The expression of molecular markers in the *her*-MO-cell mosaic embryos revealed that oscillation of the segmentation clock is specifically affected on the transplanted side (Fig. 2d, f), whereas the level of FGF signalling in the PSM, which also regulates segment positions^{18,19}, is unchanged (Supplementary Notes S2). *her1* expression is maintained in a segmental pattern on both sides of these transplanted embryos, but becomes out of phase such that the oscillation of the transplanted side advances relative to that of the control side (Fig. 2d–g and Supplementary Notes S2). This result indicates that the oscillation phase is accelerated by *her*-MO cells. Consistently, when donor cells are localized either laterally or medially in the PSM, the shape of the *her1* stripes becomes flattened

¹Department of Biological Sciences, Graduate School of Science, University of Tokyo, Hongo 7-3-1, Tokyo 113-0033, Japan. ²Division of Biological Science, Graduate School of Science, Nagoya University, Furo-cho, Nagoya 464-8602, Japan.

*These authors contributed equally to this work.

(Fig. 2d) or medially tilted (Fig. 2f), supporting the idea that *her*-MO cells locally accelerate *her1* oscillation in adjacent cells. Notably, expression of *her1* is already affected in the posterior-most synchronized oscillation zone (bracket in Fig. 2f) as well as in the travelling wave zone (intermediate and anterior PSM), indicating that the interaction between normally oscillating host and *her*-MO donor cells occurs in the posterior PSM. Indeed, high-resolution *in situ* hybridization (ISH) clearly demonstrates that the timing of *her1* transcription is locally advanced in cells that surround the transplanted *her*-MO cells in the posterior PSM (Fig. 2h, i). Furthermore, the intercellular communication that accelerates *her1* oscillation is certainly active in the synchronized oscillation zone, because the same results were obtained when *her*-MO cells were directly transplanted into the posterior PSM of normal host: segment positions and *her1* expression were locally shifted around the transplanted cells (Supplementary Notes S2). Thus, Notch-dependent intercellular communication accelerates the phase of neighbouring oscillators *in vivo*.

The 'phase acceleration' that we observed in the *in vivo* experiment is a general response of a coupled oscillator system. Indeed, local acceleration of an oscillating system by actively signalling cells is reproduced in our numerical simulation (for details, see Box 1, Supplementary Notes S3 and Supplementary Movie S1).

We conclude that the segmentation clock behaves as a coupled oscillator, in which cellular oscillators are interconnected through oscillator-linked Notch signalling.

Fluctuations in individual oscillators

In the segmentation clock, the period and phase of individual

oscillators might fluctuate for the following reasons. Rhythm-generating reactions such as transcription, translation and translocation of key factors are in themselves stochastic processes at the molecular level^{20–22}. Furthermore, the dynamics of these processes can be affected by mitosis, during which the transcription and translation of most genes are arrested²³. Changes in cell shape during cell division may also alter the efficiency of contact-dependent intercellular communication that could potentially affect the oscillation period. Indeed, we have observed active cell proliferation in the synchronized oscillation zone (Fig. 3a). Time-lapse analysis of embryos expressing a histone 2B–green fluorescent protein (GFP) fusion gene revealed that during one cycle of oscillation, 10–15% of cells experience the mitotic phase (M phase), which lasts at least 15 min (Fig. 3b, Supplementary Table S1 and Supplementary Movie S2). Moreover, given the remarkably short period of *her1* oscillation (30 min), the presence of numerous dividing cells with the 15-min M phase might not be negligible for coherent oscillation.

To test whether the cyclic expression of segmentation genes actually fluctuates *in vivo*, we determined the phase of oscillating cells in the synchronized oscillation zone using high-resolution ISH. As described above, cells in this zone repeat the ON/OFF cycle in almost complete synchrony. However, a certain proportion of these cells were out of phase (Fig. 3c); for example, some cells prematurely initiate *her1* transcription in a transcriptionally silent population (arrowhead in Fig. 3d) and others still retain *her1* messenger RNA in the cytoplasm as their surrounding cells become negative for its presence (arrow in Fig. 3e). The presence of slightly advanced and slightly delayed cells indicates that *her1* oscillation fluctuates in individual cells, and this may be caused by stochastic gene expression. Stochastic transcription in each allele within a cell can be detected by the differential distribution of exonic and intronic probes for *her1*. As is shown in Fig. 3f, one allele is being transcribed (positive for both intronic and exonic probes; arrowhead) while transcription in

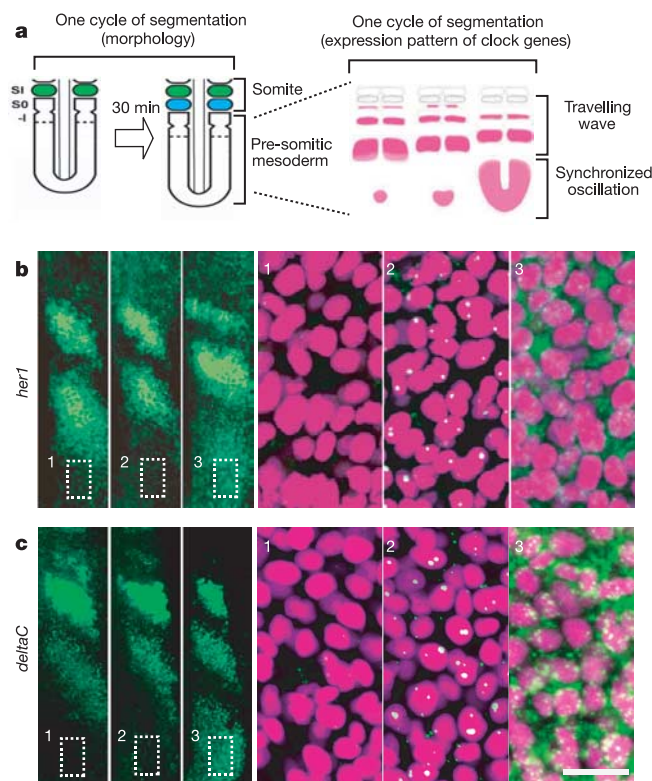


Figure 1 | Synchronized oscillation of *her1* and *deltaC* in the posterior PSM. **a**, Schematic representation of the clock oscillation during one segmentation cycle. **b**, **c**, Subcellular localization of *her1* (**b**) and *deltaC* (**c**) mRNA (green) visualized by high-resolution ISH. PSM cells display one of the following patterns of mRNA localization: no signal, nuclear dots or cytoplasmic localization, representing states with no transcripts, with active transcription, or with active translation, respectively. Numbered insets are enlarged in the panels to the right; see also Supplementary Notes S1. Magenta indicates nuclei. Scale bar, 20 μ m.

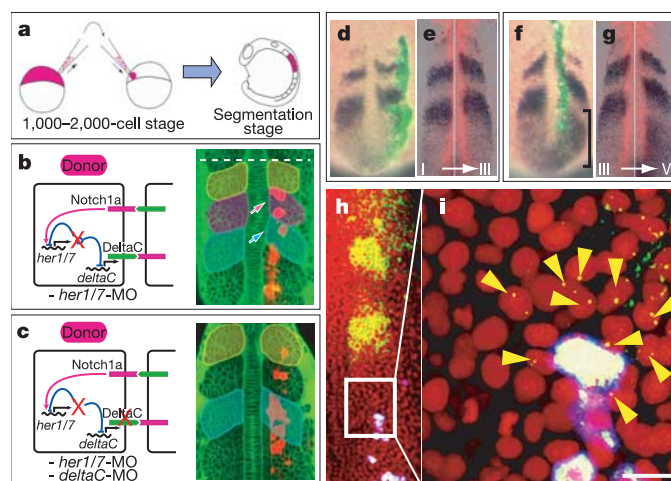


Figure 2 | Effects of actively signalling cells upon synchronized oscillation. **a**, Cell-transplantation assay at the blastula stage. **b**, Affected segment positions in a *her*-MO-cell mosaic embryo. The dashed line indicates the last normally formed segment border. **c**, Rhodamine-labelled donor cells doubly injected with *deltaC*-MO and *her*-MO do not affect segment positions. Tissue morphology is outlined by Bodipy-ceramide staining. **d**, **f**, *her1* expression (purple) in mosaic embryos that received *her*-MO cells (green). **e**, **g**, Composite images of normal *her1* expression patterns (purple) arranged so as to represent **d** and **f**. Note that the pattern on the right side is advanced compared with that on the left (see also Supplementary Notes S2). **h**, **i**, The timing of *her1* transcription is locally advanced in the area near to the explants in the posterior PSM. Nuclei, *her1* mRNA and explants are stained red, green and blue, respectively. Arrowheads indicate nuclear *her1*-positive cells. Scale bar, 20 μ m.

another allele is just terminated (positive only for exonic probes; arrow). This contrasts with the nucleus, which shows active transcription in both alleles (asterisk in Fig. 3f). Furthermore, some asynchronous cells can be found in pairs (arrows in Fig. 3g). This raises the possibility that they are sibling cells generated by a recent cell division. In support of this idea, we did not detect *her1* transcripts in either condensed or segregating chromatin even though the surrounding cells were synchronously initiating *her1* transcription (Fig. 3h, i). These data indicate that internal noise caused by both stochastic gene expression and cell division affects the oscillation period of unit oscillators, as is often seen for other biological oscillators^{24,25}.

Coupling-assisted coherent oscillation against noise

The effects of noise must be minimized to maintain coherent oscillation, and it is thought that intercellular coupling has a role in this process, as is the case for other 'noisy' biological systems^{1,2,26}. To understand the importance of Notch signalling for coherent oscillation, we cultured embryos in the presence of the γ -secretase inhibitor DAPT to reduce Notch activity. Wild-type embryos at the 6-somite stage were transiently treated with DAPT or DMSO for the

next four cycles of oscillation, and subjected to high-resolution ISH. This treatment greatly affected *her1* synchrony in the posterior PSM, although the overall expression patterns were less affected under these experimental conditions (Fig. 4a, b). In DAPT-treated embryos, the cellular localization of *her1* mRNA became variable from cell to cell (Supplementary Fig. S10) and the proportion of nuclear *her1*-positive cells was approximately 20–30% at every phase of oscillation, whereas that of the control embryos shifted from 10% to 90%, depending on the phase. Similar results were obtained by our numerical simulation in which 25 PSM cells oscillate in the presence of noise with or without intercellular coupling (Box 1 and Supplementary Notes S3). Thus, *in vivo* and *in silico* experiments demonstrate the presence of the Notch-dependent phase synchronization mechanism in the segmentation clock.

The phase synchronization was directly tested by a cell-transplantation experiment. To juxtapose two independently oscillating populations, we isolated a group of cells from the posterior PSM of a wild-type donor and directly transplanted them into another wild-type host at the same axial level (Fig. 4c). After several rounds of oscillation, the oscillation phase of the explants, some of which must have been originally out of phase, became synchronized

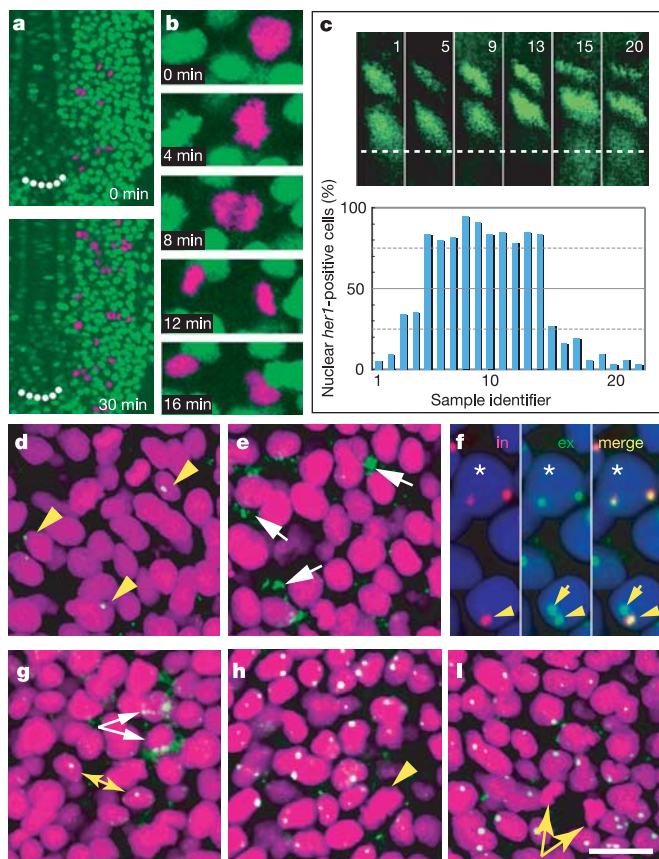


Figure 3 | Fluctuated oscillation of *her1*. **a, b**, Mitosis in the posterior PSM. Nuclei (histone 2B–GFP) that experienced mitosis during one cycle of oscillation are labelled magenta (pseudo-coloured). White dots indicate the posterior end of the axial mesoderm. **c**, The percentage of nuclear *her1*-positive cells in 22 specimens of posterior PSM (posterior to the dashed line). Specimen numbers are indicated. **d, e**, A few of the nuclear *her1*-positive (arrowheads in **d**) or cytoplasmic *her1*-positive cells (arrows in **e**) are detectable even in the OFF-state PSM. **f**, Nuclear localization of premature and fully matured *her1* transcripts, detected by intronic (in, red) and exonic (ex, green) probes, respectively. **g**, Pairs of asynchronous cells. **h, i**, Absence of *her1* transcripts in either condensed or segregating chromatin. All *her1* mRNA is detected by exonic probes except for **f**. Anterior is to the top. Scale bar, 20 μ m.

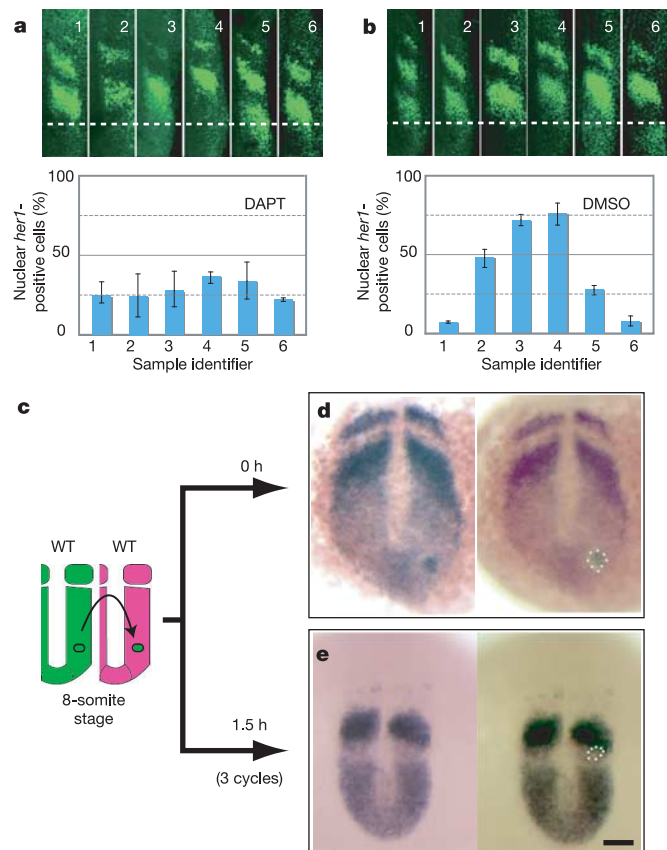


Figure 4 | Coupling-assisted phase synchronization. **a, b**, Synchrony in 10-somite-stage embryos transiently treated with DAPT (**a**) or DMSO (**b**) for 2 h, starting at the 6-somite stage. Representative expression patterns of *her1* and its synchrony (posterior to the dashed line) are shown. Error bars indicate the range of scores taken from three different z-positions in one PSM sample. **c–e**, Synchronization of out-of-phase explants. **c**, Wild-type cells of the posterior PSM were directly transplanted into normal embryos at the same axial level. **d, e**, *her1* expression (purple in the left and right panels) in these embryos just after transplantation (**d**) or 1.5 h later (**e**) is shown. At the time of transplantation, 9 out of 22 explants exhibit an oscillation phase that is different from that of the host, whereas 24 out of 26 are in phase after three rounds of oscillation. The position of biotin-labelled donor cells is detected as green staining (encircled with white dots in the right panels). Scale bar, 50 μ m.

with that of the host (Fig. 4d, e; see also Supplementary Notes S3 and Supplementary Movie S3 for simulation), indicating that phase synchronization actually works in the segmentation oscillator.

We conclude that Notch-dependent intercellular coupling is crucial for the maintenance of synchronized oscillation, by reducing the effects of internal noise in the PSM.

Discussion

By using a simple model with a minimum number of components, we have experimentally shown that the segmentation clock behaves as a typical coupled oscillator that displays strong robustness to developmental noise. However, an *in vivo* oscillation mechanism must be much more complicated than we have assumed here²⁷. We observed that simple overexpression of DeltaC and DeltaD proteins (either or both) is not sufficient for PSM explants to exert the segment-shift activity (data not shown), suggesting that an as-yet-unidentified factor(s) functions downstream of Her1/7 in parallel with Delta proteins. Fringe (a modulator of Notch signalling^{6,28,29}) or the Wnt pathway³⁰, both of which are essential in other vertebrate species, might be such factors. However, the minimum model that we used was very informative in understanding the system-level properties of the clock. Recently, several groups demonstrated that simplification of the dynamics of a few representative elements, although neglecting molecular details, can provide valuable information about

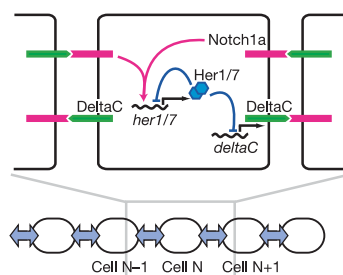
the behaviour of biological systems^{31–33}. The present study could be one such case.

It is noteworthy that vertebrate embryos use a robust system for developmental pattern formation, which has a primary function to establish a reproducible pattern in a proliferating population of cells. In this context, one may recall the classical concept of developmental biology, canalization, in which living organisms produce a standard phenotype with surprisingly few variations under different environmental and genetic conditions³⁴. This has raised the important issue of how such a developmental programme can be so resistant to various perturbations^{35,36}. This phenomenon has been partially explained by the presence of redundant elements and by self-balancing feedback regulation, which mainly functions at the level of local circuit^{37–39}. At the same time, it has been theoretically suggested that higher-order mechanisms, driven by system-level dynamics, could also ensure the stability of the genetic programme, as exemplified by the wide ranging tolerance of variation of the initial conditions during the segmentation programme in *Drosophila*^{40,41}. In our current study, we demonstrate that such higher-order mechanisms—phase synchronization and mutual entrainment—ensure the coherent oscillation of the segmentation clock, even in the presence of proliferating cells. This work therefore presents a novel molecular explanation for the mechanisms underlying developmental canalization.

Box 1 | Modelling of Notch-coupled oscillators

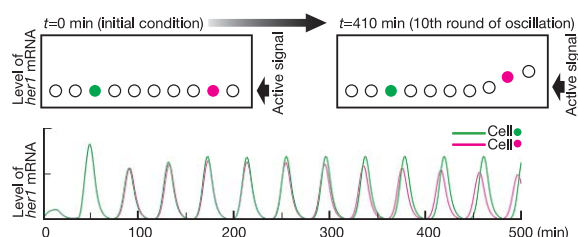
To assess the results obtained in the *in vivo* experiments, we numerically examined the phase dynamics of the segmentation clock *in silico*. The oscillation dynamics shown in the following simulation (see below) is common among coupled oscillator systems and does not depend on the driving mechanism of oscillation. Here, we adopt the 'delay model' (ref. 16), simply because the model has been most supported by molecular and genetic experiments of the segmentation clock⁴⁴ (see also Supplementary Notes S1 and S3). In this model, cell-autonomous oscillation is explained by the dynamics of mRNA and protein concentration of the transcriptional repressors *her1* and *her7*, incorporating the 'delay time' for their maturation.

We simulated oscillation in a one-dimensional array of virtual PSM cells coupled by Notch-Delta communication, as shown in Box 1 Fig. 1. Delta protein expressed in one cell is evenly distributed on both sides. Thus, the Notch signal in one cell is activated by Delta protein expressed in neighbouring cells on both sides. (For details of the simulation, see Supplementary Notes S3.)



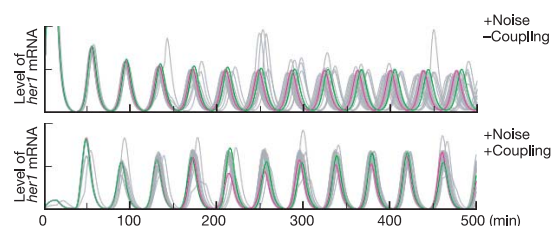
Box 1 Figure 1 | Molecular network of the segmentation clock.

To simulate the phase-shift effect induced by *her*-MO cells, we put a cell that constantly expresses high levels of Delta protein at the right end of the one-dimensional array of ten cells (Box 1 Fig. 2). The active signal from the transplanted cell influences the adjacent cell to accelerate the oscillating phase (13.4% of the oscillation phase after 500 min; see also Supplementary Fig. S7). This effect is transmitted in succession, and results in the phase shift of relatively distant cells, although the effect is still locally limited. (For details, see Supplementary Notes S3 and Supplementary Movie S1.)



Box 1 Figure 2 | Phase acceleration by activated Notch signalling.

Next, we tested the importance of intercellular coupling in the presence of noise. Based on the experimental observations (Fig. 3), the effect of noise is introduced as follows. Effect 1 is distributed frequency in each oscillator. Here, delay time for all components in each cell is distributed with the standard deviation value of 0.15 and is fixed throughout the simulation. Effect 2 is phase shift associated with mitosis. Here, dividing cells are randomly selected such that 7% of cells in a total population experience mitosis during one cycle of oscillation. During the division (supposed to last for 15 min), the delay time is doubled and the synthesis rate of mRNAs and proteins is set to half the value. At time 0, all the cells in the array oscillate synchronously. Without the coupling (up), the phase of oscillation soon becomes random. However, with coupling (down), the cell array can maintain the coherent oscillation (Box 1 Fig. 3). (For details, see Supplementary Notes S3.)



Box 1 Figure 3 | Oscillation pattern in a 25-cell array.

METHODS

High-resolution ISH and image analysis. To preserve the cellular localization of mRNAs, embryos were fixed with 4% PFA for 2 h at room temperature. During the course of the study, we found that ice-cold fixative never produced nuclear signals. DIG- or FITC-labelled cRNA probes were detected with TSA-Alexa488 (Molecular Probes: T-20932) or TSA-Alexa647 (Molecular Probes: T-20936), according to the manufacturer's instructions and ref. 42. Detailed procedure is also described in ref. 43. To determine the numbers of nuclear *her1*-positive cells, whole PSMs of flat-mounted embryos were optically scanned at 1- μ m intervals (Olympus: FV500). PSM cells located at the levels between the 5th and 15th adaxial cells from the posterior end were subjected to this analysis. Three distinct z-positions in each PSM were sterically analysed using Volocity (Improvision). The values from each assessment were analysed graphically using Microsoft Excel.

Time-lapse imaging of mitosis. To visualize mitosis in the posterior PSM, embryos were injected with histone 2B-EGFP mRNA (100 ng μ l⁻¹) at the one-cell stage. Embryos mounted in an LMP-agarose gel were time-lapse imaged with a confocal microscope (Olympus: FV500) for 30 min. Each PSM was scanned with six optical slices (10- μ m thickness) at 1-min intervals and z-stack images at each time point were processed for movies using Photoshop and ImageReady (Adobe).

Detailed description of other methods including cell transplantation and DAPT treatment are provided in Supplementary Methods.

Received 26 October 2005; accepted 5 May 2006.

- Winfree, A. T. Biological rhythms and the behavior of populations of coupled oscillators. *J. Theor. Biol.* **16**, 15–42 (1967).
- Winfree, A. T. *The Geometry of Biological Time* (Springer, New York, 2000).
- Palmerim, I., Henrique, D., Ish-Horowicz, D. & Pourquié, O. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* **91**, 639–648 (1997).
- Pourquié, O. The segmentation clock: converting embryonic time into spatial pattern. *Science* **301**, 328–330 (2003).
- Saga, Y. & Takeda, H. The making of the somite: molecular events in vertebrate segmentation. *Nature Rev. Genet.* **2**, 835–845 (2001).
- Dale, J. K. *et al.* Periodic notch inhibition by *lunatic fringe* underlies the chick segmentation clock. *Nature* **421**, 275–278 (2003).
- Cooke, J. Control of somite number during morphogenesis of a vertebrate, *Xenopus laevis*. *Nature* **254**, 196–199 (1975).
- Cooke, J. & Zeeman, E. C. A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *J. Theor. Biol.* **58**, 455–476 (1976).
- Cooke, J. A gene that resuscitates a theory—somitogenesis and a molecular oscillator. *Trends Genet.* **14**, 85–88 (1998).
- Holley, S. A., Julich, D., Rauch, G. J., Geisler, R. & Nusslein-Volhard, C. *her1* and the notch pathway function within the oscillator mechanism that regulates zebrafish somitogenesis. *Development* **129**, 1175–1183 (2002).
- Hirata, H. *et al.* Oscillatory expression of the bHLH factor *Hes1* regulated by a negative feedback loop. *Science* **298**, 840–843 (2002).
- Bessho, Y., Hirata, H., Masamizu, Y. & Kageyama, R. Periodic repression by the bHLH factor *Hes7* is an essential mechanism for the somite segmentation clock. *Genes Dev.* **17**, 1451–1456 (2003).
- Henry, C. A. *et al.* Two linked hairy/Enhancer of split-related zebrafish genes, *her1* and *her7*, function together to refine alternating somite boundaries. *Development* **129**, 3693–3704 (2002).
- Oates, A. C. & Ho, R. K. Hairy/E(spl)-related (Her) genes are central components of the segmentation oscillator and display redundancy with the Delta/Notch signaling pathway in the formation of anterior segmental boundaries in the zebrafish. *Development* **129**, 2929–2946 (2002).
- Jiang, Y. J. *et al.* Notch signalling and the synchronization of the somite segmentation clock. *Nature* **408**, 475–479 (2000).
- Lewis, J. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. *Curr. Biol.* **13**, 1398–1408 (2003).
- Holley, S. A., Geisler, R. & Nusslein-Volhard, C. Control of *her1* expression during zebrafish somitogenesis by a delta-dependent oscillator and an independent wave-front activity. *Genes Dev.* **14**, 1678–1690 (2000).
- Sawada, A. *et al.* Fgf/MAPK signalling is a crucial positional cue in somite boundary formation. *Development* **128**, 4873–4880 (2001).
- Dubrule, J., McGrew, M. J. & Pourquié, O. FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* **106**, 219–232 (2001).
- McAdams, H. H. & Arkin, A. Stochastic mechanisms in gene expression. *Proc. Natl Acad. Sci. USA* **94**, 814–819 (1997).
- Shav-Tal, Y. *et al.* Dynamics of single mRNPs in nuclei of living cells. *Science* **304**, 1797–1800 (2004).
- Blake, W. J., Kærn, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003).
- Prescott, D. M. & Bender, M. A. Synthesis of RNA and protein during mitosis in mammalian tissue culture cells. *Exp. Cell Res.* **26**, 260–268 (1962).
- Nagoshi, E. *et al.* Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells. *Cell* **119**, 693–705 (2004).
- Welsh, D. K., Yoo, S. H., Liu, A. C., Takahashi, J. S. & Kay, S. A. Bioluminescence imaging of individual fibroblasts reveals persistent, independently phased circadian rhythms of clock gene expression. *Curr. Biol.* **14**, 2289–2295 (2004).
- Ueda, H. R., Hirose, K. & Iino, M. Intercellular coupling mechanism for synchronized and noise-resistant circadian oscillators. *J. Theor. Biol.* **216**, 501–512 (2002).
- Rida, P. C., Le Minh, N. & Jiang, Y. J. A Notch feeling of somite segmentation and beyond. *Dev. Biol.* **265**, 2–22 (2004).
- Evrard, Y. A., Lun, Y., Aulehla, A., Gan, L. & Johnson, R. L. *lunatic fringe* is an essential mediator of somite segmentation and patterning. *Nature* **394**, 377–381 (1998).
- Zhang, N. & Gridley, T. Defects in somite formation in *lunatic fringe*-deficient mice. *Nature* **394**, 374–377 (1998).
- Aulehla, A. *et al.* Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev. Cell* **4**, 395–406 (2003).
- Ma'ayan, A. *et al.* Formation of regulatory patterns during signal propagation in a mammalian cellular network. *Science* **309**, 1078–1083 (2005).
- Brandman, O., Ferrell, J. E. Jr, Li, R. & Meyer, T. Interlinked fast and slow positive feedback loops drive reliable cell decisions. *Science* **310**, 496–498 (2005).
- Bornholdt, S. Systems biology. Less is more in modeling large genetic networks. *Science* **310**, 449–451 (2005).
- Waddington, C. Canalization of development and the inheritance of acquired characters. *Nature* **150**, 563–565 (1942).
- Houchmandzadeh, B., Wieschaus, E. & Leibler, S. Establishment of developmental precision and proportions in the early *Drosophila* embryo. *Nature* **415**, 798–802 (2002).
- Lucchetta, E. M., Lee, J. H., Fu, L. A., Patel, N. H. & Ismagilov, R. F. Dynamics of *Drosophila* embryonic patterning network perturbed in space and time using microfluidics. *Nature* **434**, 1134–1138 (2005).
- Freeman, M. Feedback control of intercellular signalling in development. *Nature* **408**, 313–319 (2000).
- Eldar, A. *et al.* Robustness of the BMP morphogen gradient in *Drosophila* embryonic patterning. *Nature* **419**, 304–308 (2002).
- Hall, B. K. & Olson, W. M. *Keywords and Concepts in Evolutionary Developmental Biology* (Harvard Univ. Press, Cambridge, Massachusetts, 2003).
- von Dassow, G., Meir, E., Munro, E. M. & Odell, G. M. The segment polarity network is a robust developmental module. *Nature* **406**, 188–192 (2000).
- Ingolia, N. T. Topology and robustness in the *Drosophila* segment polarity network. *PLoS Biol.* **2**, e123 (2004).
- Kosman, D. *et al.* Multiplex detection of RNA expression in *Drosophila* embryos. *Science* **305**, 846 (2004).
- Julich, D. *et al.* *beamter/deltaC* and the role of Notch ligands in the zebrafish somite segmentation, hindbrain neurogenesis and hypochord differentiation. *Dev. Biol.* **286**, 391–404 (2005).
- Hirata, H. *et al.* Instability of *Hes7* protein is crucial for the somite segmentation clock. *Nature Genet.* **36**, 750–754 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Fujimori for providing human histone 2B-EGFP cDNA; D. Tautz for the technical advice regarding the synthesis of intronic probes for *her1*; T. Tomita and T. Iwatsubo for their gift of DAPT; and Y. J. Jiang for Notch1a mutant (*des^{th35b}*). We also thank H. Ueda, T. Uemura, M. Takeichi, A. Takamatsu and J. Lewis for their advice and for critical reading of the manuscript. This work was supported by a postdoctoral fellowship and in part by Grants-in-Aid for Scientific Research Priority Area Genome Science and Organized Research Combination System, both of which are from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.H. (kazu@biol.s.u-tokyo.ac.jp) or H.T. (htakeda@biol.s.u-tokyo.ac.jp).

LETTERS

Stabilization of the disk around β Pictoris by extremely carbon-rich gas

Aki Roberge¹, Paul D. Feldman², Alycia J. Weinberger³, Magali Deleuil⁴ & Jean-Claude Bouret⁴

The edge-on disk surrounding the nearby young star β Pictoris is the archetype of 'debris disks', which are composed of dust and gas produced by collisions between—and evaporation of—planetesimals, analogues of Solar System comets and asteroids. These disks may provide insight into the formation and early evolution of terrestrial planets. Previous work on β Pic concluded that the disk gas has roughly solar abundances of elements¹, but this poses a problem because such gas should rapidly be blown away from the star, contrary to observations showing a stable gas disk in keplerian rotation^{1,2}. Here we report the detection of singly and doubly ionized carbon (C II, C III) and neutral atomic oxygen (O I) gas in the β Pic disk. Carbon is extremely overabundant relative to every other measured element. This appears to solve the problem of the stable gas disk, because the carbon overabundance should keep the gas disk in keplerian rotation³. The overabundance may indicate that the gas is produced from material more carbon-rich than expected of Solar System analogues.

β Pic is an A5 V star, indicating that it is approximately 1.8 times more massive than the Sun. It appears to have solar elemental abundances⁴ and is 8–20 million years old⁵. Narrow atomic absorption lines were studied in the spectra of β Pic even before the star was known to have a dust disk⁶. The absorption features fall into one of two categories. Almost every line observed shows an unvarying narrow absorption at the radial velocity of the star. This component contains the bulk of the circumstellar gas and is called the stable component¹. On the wings of most absorption lines are broad, variable absorption features, which are typically redshifted with respect to the star. These features arise from gas falling towards the star at high velocity, probably produced by vaporization of star-grazing planetesimals⁷.

Various studies of the dynamics of the β Pic stable gas have shown that there is a serious gap in our understanding. The force of stellar radiation pressure on many of the observed atomic and ionic species should rapidly blow them away from the star. The very existence of gas at the velocity of the star (the bulk of the β Pic circumstellar gas) is puzzling. Workers were forced to postulate the existence of an undetected, dense hydrogen torus to brake some species through gas drag¹. The problem got worse after the detection of resonantly scattered emission from gaseous species in the β Pic disk^{2,8}. These spatially resolved spectra show a gas disk in keplerian rotation between 13 astronomical units (AU) and a few hundred AU from the star. Analysis of the gas dynamics demonstrated that the some of the observed species should not be in keplerian rotation, unless the total gas mass is much greater than the observed gas mass².

We used spectra from the Far Ultraviolet Spectroscopic Explorer (FUSE) to measure the volatile species C II, C III, and most importantly, O I, in the circumstellar disk. These species are seen in absorption against broad stellar emission lines (Fig. 1). Models of

the absorption lines were compared to the data and the best absorption parameters (Table 1) determined through χ^2 minimization (details of the observations, models, and analysis appear in the Supplementary Information). Combined with a previous measurement of C I gas⁹, the

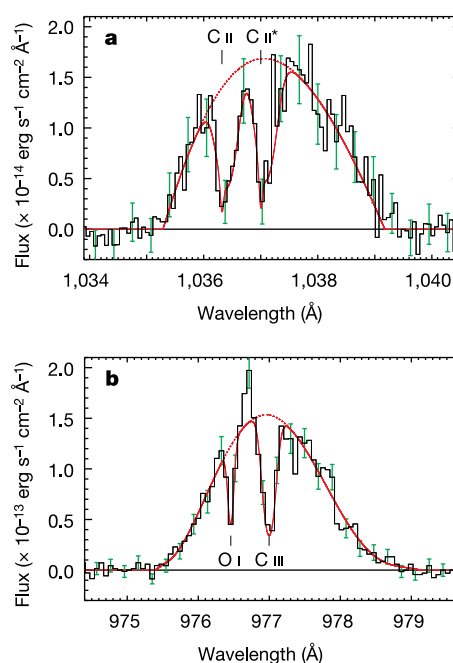


Figure 1 | Circumstellar absorption lines in far-ultraviolet spectra of β Pic. β Pic was observed by FUSE once in 2002 and twice in 2001; the higher signal-to-noise spectra from 2002 are shown here. The error bars (shown in green) are the standard deviations of the flux values ($\pm 1\sigma$). Details of the observations and analysis appear in the online Supplementary Information. Atomic absorption lines from C II, O I, and C III are superimposed on broad emission lines arising in the chromosphere and transition zone of the star¹⁷. We model each emission feature with a polynomial (dashed red line). The total model (emission model times best absorption model from χ^2 minimization) is overplotted with a solid red line. **a**, The C II absorption doublet at wavelengths $\lambda = 1036.3$ Å and $\lambda = 1037.0$ Å superimposed on the O VI $\lambda = 1038$ Å emission line. There is also some C II emission blended with the blue wing of the O VI line¹⁸. The asymmetric shape of the C II absorption lines indicates there are two absorption components at different velocities. **b**, The O I $\lambda = 976.5$ Å and C III $\lambda = 977.0$ Å absorption lines superimposed on the C III $\lambda = 977$ Å emission line. These lines are fitted with single absorption components. Underestimation of the O I column density due to saturation of the absorption line is unlikely, because the line is not strongly saturated. The interstellar contribution to the O I absorption line is negligible (see Supplementary Information).

¹Exoplanets and Stellar Astrophysics Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA. ²Department of Physics & Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA. ³Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington DC 20015, USA. ⁴Laboratoire d'Astrophysique de Marseille, BP 8, F-13376 Marseille cedex 12, France.

Table 1 | Characteristics of the C II, O I, and C III absorption lines

Species	Wavelength (Å)	ν_{FUSE} (km s ⁻¹)	$\nu_{\text{heliocentric}}$ (km s ⁻¹)	b (km s ⁻¹)	N (atoms cm ⁻²)
C II	1036.34	$-5.0^{+5.8}_{-11.5}$	20	$1.0^{+4.7}_{-1.0}$	$1.0^{+1.7}_{-0.3} \times 10^{16}$
		$20.0^{+5.8}_{-9.3}$	45	$61.0^{+11.8}_{-6.8}$	$4.0^{+6.7}_{-1.0} \times 10^{14}$
C II*	1037.02	$-5.0^{+5.8}_{-11.5}$	20	$1.0^{+4.7}_{-1.0}$	$1.0^{+1.2}_{-0.3} \times 10^{16}$
		$20.0^{+5.8}_{-9.3}$	45	$61.0^{+11.8}_{-6.8}$	$4.0^{+4.9}_{-1.2} \times 10^{14}$
O I	976.45	$3.0^{+3.2}_{-2.8}$	20	$14.0^{+3.6}_{-2.6}$	$3.6^{+0.8}_{-0.6} \times 10^{15}$
O I*	977.96	...	...	...	$\leq 2.6 \times 10^{15}$
O I**	978.62	...	...	...	$\leq 8.7 \times 10^{14}$
C III	977.02	$-7.0^{+2.1}_{-3.2}$	10	$32.0^{+4.1}_{-3.6}$	$4.5^{+0.2}_{-0.6} \times 10^{13}$

This table shows the best absorption parameters obtained through χ^2 minimization; the models were compared to the 2002 FUSE spectra. Details of the models and fitting procedures appear in the Supplementary Information. The errors given are the standard deviations of the parameters ($\pm 1\sigma$) arising from statistical errors. In column 1, * or ** indicates that the line arises from the first or second excited fine-structure energy level of the atom; otherwise, the line arises from the ground energy level. Column 2 shows the vacuum rest wavelengths of the absorption lines. In column 3, ν_{FUSE} is the velocity of the gas in the FUSE wavelength scale. FUSE does not carry an onboard calibration lamp, so the absolute wavelength scales of the spectra may have zero-point offsets. The accuracy of the relative wavelength calibration is about 8 km s⁻¹. Column 4 shows the heliocentric velocity of the gas, obtained by shifting the stable gas components to the stellar heliocentric velocity (20 km s⁻¹). Column 5 shows the best Doppler broadening parameters of the absorption components, which is a measure of the intrinsic width of the component. The last column shows the column densities in the various absorption components. The upper limits given for O I* and O I** were calculated assuming statistical population of the oxygen fine-structure levels (see Supplementary Information S2.2). The smallest possible value for the total O I column density is the measured ground level O I column density (row 5). The largest possible value is the sum of the measured O I column density and the upper limits on the O I* and O I** column densities (rows 5, 6, and 7).

C II measurement completes the inventory of stable carbon gas in the disk. O I is the only ionization state of oxygen observed to date, and the only one likely to be abundant, given the high first ionization energy of oxygen. We also used archival Hubble Space Telescope Space Telescope Imaging Spectrograph (STIS) high-resolution far-ultraviolet spectra to search for absorption lines from a few species that have not yet been observed (Cr I, Ni I, P I and P II), but none were detected.

Previous observations of the β Pic gas were primarily of metallic species¹. The addition of our new measurements provides the most complete inventory of the gas in any debris disk. The β Pic elemental abundances, found by summing the abundances of the various ionization stages of each element, are shown in Table 2, along with the elemental abundances for the Sun¹⁰, a carbonaceous chondrite meteorite¹⁰, and dust from comet Halley¹¹. Table 2 uses results from

studies of the β Pic gas stretching over more than ten years^{1,2,9,12–14} (see Supplementary Information). A plot of the elemental abundances from Table 2 appears in Fig. 2. It shows that the composition of the β Pic gas is dissimilar to the composition of all three comparison bodies. In particular, carbon is extremely overabundant relative to every other measured element. The β Pic C/O and C/Fe ratios are 18 and 16 times the solar value, respectively. The lithophile elements (for example, Mg, Al) have roughly solar abundance relative to each other, as do the siderophile elements (for example, Fe, Ni). However, the lithophile elements as a group seem slightly underabundant relative to the siderophile elements.

Neither oxygen nor carbon feels strong radiation pressure in the β Pic disk, since the star is relatively faint in the far-ultraviolet where the strong absorption lines of these species lie. By contrast, metals such as Na and Fe feel extremely strong radiation pressure and could be blown out of the system, producing apparent C and O overabundances relative to these elements. But carbon is overabundant relative to every other element, including ones that feel similarly weak radiation pressure, like oxygen. This leads us to suspect that the overabundance reflects the composition of the parent material, which would have to be more carbon-rich than the obvious Solar System analogue materials. The bulk composition of the β Pic planetesimals might be carbon-rich. There was a large increase in organics relative to water in the material excavated from comet Tempel 1 during its collision with the Deep Impact spacecraft¹⁵. Or the planetesimals might be selectively outgassing volatile carbonaceous compounds, trace amounts of which are still found in carbonaceous chondrite meteorites (J. Nuth, personal communication). There are probably many other possible explanations.

A new analysis of the gas dynamics in β Pic which took into account Coulomb forces between gaseous ions, neutral atoms, and charged dust showed that these interactions increase the effective collision cross-sections, dramatically reducing the need for additional unobserved braking gas³. But unless the dust grains are very highly charged, additional braking gas is still required. However, that analysis assumed solar abundances in the gas. If one considers our measured midplane abundances in conjunction with Coulomb forces, the problem may be solved.

The ratio of the radiation and gravitational forces felt by an atom is the radiation pressure coefficient (β). The β -value for a particular species depends on the brightness of the central star at the wavelengths of the absorption lines of the species. Neutral elements that

Table 2 | Elemental abundances in the β Pic stable gas

Element	β Pic column density (atoms cm ⁻²)	Sun N (EI)	CI chondrite N (EI)	Halley dust N (EI)
H	$\leq \text{few} \times 10^{19}$	2.88×10^{10}	$5.50^{+0.67}_{-0.60} \times 10^6$	$(1.09 \pm 0.21) \times 10^7$
O	$(3-8) \times 10^{15}$	$1.41^{+0.17}_{-0.15} \times 10^7$	$7.55^{+0.36}_{-0.34} \times 10^6$	$(4.81 \pm 0.59) \times 10^6$
C	$5.0^{+2.3}_{-1.1} \times 10^{16}$	$7.08^{+0.68}_{-0.62} \times 10^6$	$7.72^{+1.14}_{-1.00} \times 10^5$	$(4.40 \pm 0.89) \times 10^6$
N	...	$1.95^{+0.56}_{-0.44} \times 10^6$	$5.54^{+0.97}_{-0.82} \times 10^4$	$(2.27 \pm 0.76) \times 10^5$
Mg	$\geq 2 \times 10^{13}$	$1.00^{+0.05}_{-0.05} \times 10^6$	$1.04^{+0.05}_{-0.05} \times 10^6$	5.41×10^5
Si	1×10^{14}	$1.00^{+0.05}_{-0.05} \times 10^6$	$1.00^{+0.05}_{-0.05} \times 10^6$	$(1.00 \pm 0.10) \times 10^6$
Fe	$(3.7 \pm 0.5) \times 10^{14}$	$8.13^{+0.58}_{-0.54} \times 10^5$	$8.63^{+0.62}_{-0.58} \times 10^5$	$(2.81 \pm 0.49) \times 10^5$
S	1×10^{13}	$4.63^{+0.45}_{-0.41} \times 10^5$	$4.45^{+0.43}_{-0.39} \times 10^5$	$(3.89 \pm 1.24) \times 10^5$
Al	4.5×10^{12}	$8.51^{+0.40}_{-0.38} \times 10^4$	$8.31^{+0.39}_{-0.37} \times 10^4$	$(3.68 \pm 0.92) \times 10^4$
Ca	$2.6^{+5.3}_{-0.9} \times 10^{13}$	$6.61^{+0.47}_{-0.44} \times 10^4$	$6.00^{+0.43}_{-0.40} \times 10^4$	$(3.41 \pm 1.03) \times 10^4$
Na	$\geq 3 \times 10^{13}$	$5.75^{+0.41}_{-0.38} \times 10^4$	$5.75^{+0.41}_{-0.38} \times 10^4$	$(5.41 \pm 3.24) \times 10^4$
Ni	1.5×10^{13}	$4.78^{+0.34}_{-0.32} \times 10^4$	$4.78^{+0.34}_{-0.32} \times 10^4$	$(2.22 \pm 1.14) \times 10^4$
Cr	3.5×10^{12}	$1.26^{+0.15}_{-0.14} \times 10^4$	$1.31^{+0.16}_{-0.14} \times 10^4$	$(4.86 \pm 1.08) \times 10^3$
P	$\leq 9.3 \times 10^{13}$	$8.91^{+0.86}_{-0.78} \times 10^3$	$7.83^{+0.76}_{-0.69} \times 10^3$	...
Mn	3×10^{12}	$7.08^{+0.51}_{-0.47} \times 10^3$	$9.17^{+0.66}_{-0.61} \times 10^3$	$(2.70 \pm 1.08) \times 10^3$
Zn	$(2-3) \times 10^{11}$	$1.20^{+0.12}_{-0.11} \times 10^3$	$1.25^{+0.12}_{-0.11} \times 10^3$	...

The total abundance for each element in the β Pic stable gas (column density) was found by summing the column densities of the various ionization stages of the element, which appear in Supplementary Table 2. $N(\text{EI})$ is the elemental abundance normalized to the number of Si atoms, where $N(\text{Si}) = 10^6$. The errors given are the standard deviations of the values ($\pm 1\sigma$), where available.

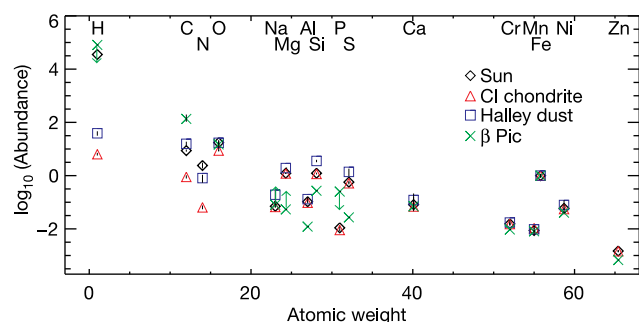


Figure 2 | Bulk composition of β Pic stable circumstellar gas. The elemental abundances for the β Pic midplane gas are shown with green crosses, the Sun¹⁰ with black diamonds, a carbonaceous chondrite meteorite¹⁰ with red triangles, and comet Halley dust¹¹ with blue squares. The y axis shows the logarithmic elemental abundance, normalized to Fe. The vertical lines on top of the data points show the standard deviations of the abundances ($\pm 1\sigma$), where available. Upper and lower limits are indicated with arrows. Details of the measurements used in constructing this plot appear in the Supplementary Information.

feel strong radiation pressure are quickly ionized before they can be blown away from the star; therefore, the species that must be braked are ionized. Coulomb interactions between the ions couple them together into a single fluid³. If the effective radiation pressure coefficient of the fluid (β_{eff}) is < 0.5 , it will be bound to the star and the gas will be self-braking. Because carbon is abundant and moderately ionized in the β Pic disk, it is an important constituent of the ionic fluid. β Pic is faint in the far-ultraviolet, where the strong C I and C II absorption lines lie, so carbon feels negligible radiation pressure in the β Pic disk. Increasing the carbon abundance lowers β_{eff} . The ionic fluid will be self-braking ($\beta_{\text{eff}} < 0.5$) if the carbon abundance is enhanced by a factor of ten or more over the solar abundance³. Carbon in the β Pic stable gas is enhanced relative to oxygen by a factor of 18 over the solar C/O abundance, so additional unseen gas or highly charged grains are not needed to keep the gas disk in keplerian rotation.

Why the gas composition is just what is required for keplerian rotation is unknown. Answering this question will require similar observations and analyses to be done for many debris disks. However, FUSE spectra of another debris disk may provide a hint¹⁶. Circumstellar gas and dust have been detected around the B9 primary star of the σ Herculis binary system. In this case, the gas is moving away from the star. It has been suggested that the gas and dust are produced by collisions among planetesimals located about 20 AU from the star, where the orbits of small bodies become unstable in the binary system. A radiatively driven wind is generated as the high radiation pressure from this ultraviolet-bright star blows away the circumstellar gas. This might be the more common fate of debris gas around a high mass star, and perhaps would have occurred in the β Pic disk if

not for the overabundance of carbon in its circumstellar gas. Of course, this does not explain how the carbon overabundance occurred in the first place. A great deal more observational and theoretical work is required to determine whether the gas composition reflects the bulk composition of the parent bodies and then determine how such carbon-rich planetesimals could have formed in the β Pic protoplanetary disk.

Received 23 January; accepted 24 April 2006.

1. Lagrange, A.-M. *et al.* The β Pictoris circumstellar disk. XXIV. Clues to the origin of the stable gas. *Astron. Astrophys.* **330**, 1091–1108 (1998).
2. Brandeker, A., Liseau, R., Olofsson, G. & Fridlund, M. The spatial structure of the β Pictoris gas disk. *Astron. Astrophys.* **413**, 681–691 (2004).
3. Fernández, R., Brandeker, A. & Wu, Y. Braking the gas in the β Pictoris disk. *Astrophys. J.* (in the press); preprint at (<http://arxiv.org/abs/astro-ph/0601244>) (2006).
4. Holweger, H., Hempel, M., van Thiel, T. & Kaufer, A. The surface composition of β Pictoris. *Astron. Astrophys.* **320**, L49–L52 (1997).
5. Zuckerman, B., Song, I., Bessell, M. S. & Webb, R. A. The β Pictoris moving group. *Astrophys. J.* **562**, L87–L90 (2001).
6. Slettebak, A. Some interesting bright southern stars of early type. *Astrophys. J.* **197**, 137–138 (1975).
7. Beust, H., Vidal-Madjar, A., Ferlet, R. & Lagrange-Henri, A. M. The β Pictoris circumstellar disk. X—Numerical simulations of infalling evaporating bodies. *Astron. Astrophys.* **236**, 202–216 (1990).
8. Olofsson, G., Liseau, R. & Brandeker, A. Widespread atomic gas emission reveals the rotation of the β Pictoris disk. *Astrophys. J.* **563**, L77–L80 (2001).
9. Roberge, A. *et al.* High-resolution Hubble Space Telescope STIS spectra of C I and CO in the β Pictoris circumstellar disk. *Astrophys. J.* **538**, 904–910 (2000).
10. Lodders, K. Solar System abundances and condensation temperatures of the elements. *Astrophys. J.* **591**, 1220–1247 (2003).
11. Jessberger, E. K., Christoforidis, A. & Kissel, J. Aspects of the major element composition of Halley's dust. *Nature* **332**, 691–695 (1988).
12. Freudling, W., Lagrange, A.-M., Vidal-Madjar, A., Ferlet, R. & Forveille, T. Gas around β Pictoris: An upper limit on the H I content. *Astron. Astrophys.* **301**, 231–235 (1995).
13. Vidal-Madjar, A. *et al.* HST-GHRS observations of β Pictoris: Additional evidence for infalling comets. *Astron. Astrophys.* **290**, 245–258 (1994).
14. Crawford, I. A., Spyromilio, J., Barlow, M. J., Diego, F. & Lagrange, A.-M. Ultra-high-resolution observations of Ca K line variations in the β Pictoris disc. *Mon. Not. R. Astron. Soc.* **266**, L65–L68 (1994).
15. A'Hearn, M. F. *et al.* Deep Impact: excavating comet Tempel 1. *Science* **310**, 258–264 (2005).
16. Chen, C. H. & Jura, M. The low-velocity wind from the circumstellar matter around the B9V star σ Herculis. *Astrophys. J.* **582**, 443–448 (2003).
17. Deleuil, M. *et al.* Is β Pictoris an active star? *Astrophys. J.* **557**, L67–L70 (2001).
18. Bouret, J.-C. *et al.* A chromospheric scenario for the activity of β Pictoris, as revealed by FUSE. *Astron. Astrophys.* **390**, 1049–1061 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Lindler for recalibrating the archival STIS spectra used in this paper. The work at JHU was supported by NASA.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.R. (akir@milkyway.gsfc.nasa.gov).

A constitutive law for dense granular flows

Pierre Jop¹, Yoël Forterre¹ & Olivier Pouliquen¹

A continuum description of granular flows would be of considerable help in predicting natural geophysical hazards or in designing industrial processes. However, the constitutive equations for dry granular flows, which govern how the material moves under shear, are still a matter of debate^{1–10}. One difficulty is that grains can behave¹¹ like a solid (in a sand pile), a liquid (when poured from a silo) or a gas (when strongly agitated). For the two extreme regimes, constitutive equations have been proposed based on kinetic theory for collisional rapid flows¹², and soil mechanics for slow plastic flows¹³. However, the intermediate dense regime, where the granular material flows like a liquid, still lacks a unified view and has motivated many studies over the past decade¹⁴. The main characteristics of granular liquids are: a yield criterion (a critical shear stress below which flow is not possible) and a complex dependence on shear rate when flowing. In this sense, granular matter shares similarities with classical visco-plastic fluids such as Bingham fluids. Here we propose a new constitutive relation for dense granular flows, inspired by this analogy and recent numerical^{15,16} and experimental work^{17–19}. We then test our three-dimensional (3D) model through experiments on granular flows on a pile between rough sidewalls, in which a complex 3D flow pattern develops. We show that, without any fitting parameter, the model gives quantitative predictions for the flow shape and velocity profiles. Our results support the idea that a simple visco-plastic approach can quantitatively capture granular flow properties, and could serve as a basic tool for modelling more complex flows in geophysical or industrial applications.

Advances in our understanding of dense granular flows have recently been made by comparing different flow configurations¹⁴. The simplest configuration from a rheological point of view is the one sketched in the inset to Fig. 1. A granular material confined under a normal stress P in between two rough planes is sheared at a

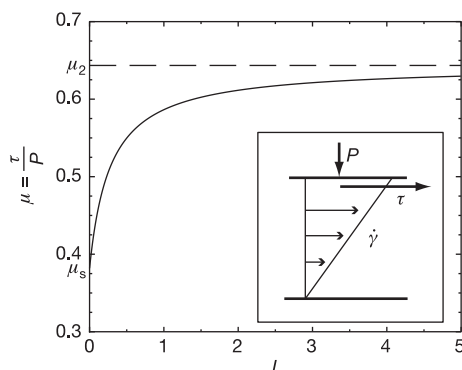


Figure 1 | Friction coefficient μ as a function of the dimensionless parameter I ($\mu_s = \tan(20.9^\circ)$, $\mu_2 = \tan(32.76^\circ)$ and $I_0 = 0.279$). Inset, definition of the pressure P , the shear stress τ , and the shear rate $\dot{\gamma}$ in the simple plane shear configuration.

given shear rate $\dot{\gamma}$ by applying a shear stress τ . In refs 15 and 16, for stiff particles the shear stress is shown, using dimensional arguments and numerical simulations, to be proportional to the normal stress, with a coefficient of proportionality that is a function of a single dimensionless number, called the inertial number I :

$$\tau = \mu(I)P \text{ with } I = \dot{\gamma}d/(P/\rho_s)^{0.5} \quad (1)$$

where $\rho(I)$ is the friction coefficient, d is the particle diameter and μ_s is the particle density. They found that the volume fraction Φ of the sample is also a function of I but varies only slightly in the dense regime. The inertial number, which is the square root of the Savage number²⁰ or of the Coulomb number²¹ introduced previously in the literature, can be interpreted as the ratio between two timescales, a macroscopic deformation timescale ($1/\dot{\gamma}$) and an inertial timescale ($d^2\rho_s/P$)^{0.5}. By confronting results from the simple shear test with experimental measurements of granular flows on rough inclined planes^{17,22}, it can be shown that the friction coefficient $\mu(I)$ has the shape given in Fig. 1. It starts from a critical value of μ_s at zero shear rate and converges to a limiting value of μ_2 at high I . The following friction law can then be proposed, compatible with the experiments¹⁹:

$$\mu(I) = \mu_s + (\mu_2 - \mu_s)/(I_0/I + 1) \quad (2)$$

where I_0 is a constant. Very recently, this simple description of granular flows has been successful in predicting two-dimensional configurations, capturing velocity profiles on inclined planes^{14,23} and important features of flows on a pile¹⁹. However, the simple scalar law (equation (1)) cannot be applied in more complex flows where shear in different

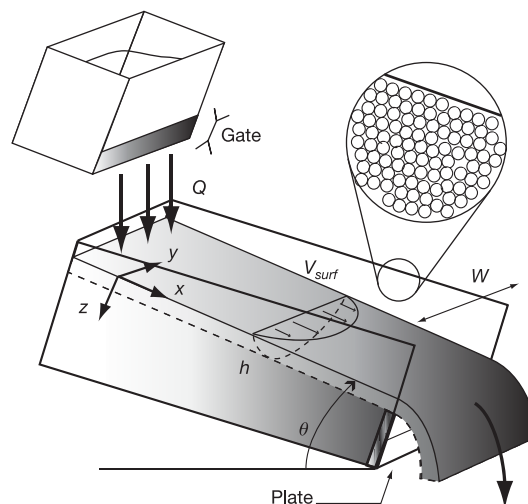


Figure 2 | Experimental set-up of granular flows on a pile between rough sidewalls. The channel is partially closed at the bottom end to create a static pile on top of which the grains flow¹⁹. The sidewalls are made rough by gluing one layer of beads on them. The channel is 120 cm long and the width W varies from 0.9 cm ($16.5d$) up to 28.9 cm ($546d$).

¹IUSTI, CNRS UMR 6595, Université de Provence, 5 rue Enrico Fermi, 13453 Marseille cedex 13, France.

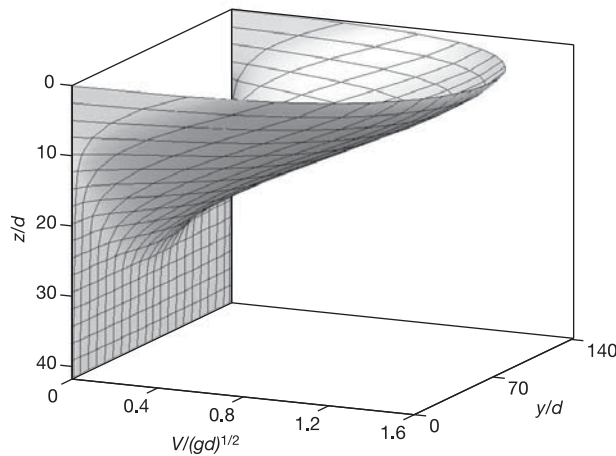


Figure 3 | Typical 3D velocity profile predicted by the rheology ($W = 142d$, $\theta = 22.6^\circ$, $Q/d^{3/2}g^{1/2} = 15.2$). For clarity only one quarter of the lines of the 71×80 computational grid is plotted.

directions is present and where a full three-dimensional rheology is needed.

We therefore propose the following 3D generalization of the friction law for a granular material. The basic assumption consists in neglecting the small variation of the volume fraction observed in the dense regime. The granular material is then described as an incompressible fluid with the internal stress tensor given by the following relations:

$$\sigma_{ij} = -P\delta_{ij} + \tau_{ij} \text{ and } \tau_{ij} = \eta(|\dot{\gamma}|, P)\dot{\gamma}_{ij} \quad (3)$$

with $\eta(|\dot{\gamma}|, P) = \mu(I)P/|\dot{\gamma}|$ and $I = |\dot{\gamma}|d/(P/\rho_s)^{0.5}$

where $\dot{\gamma}_{ij} = \partial u_i/\partial x_j + \partial u_j/\partial x_i$ is the strain rate tensor and $|\dot{\gamma}| = (0.5\dot{\gamma}_{ij}\dot{\gamma}_{ij})^{0.5}$ is the second invariant of $\dot{\gamma}_{ij}$. In this rheology, P represents an isotropic pressure, and $\eta(|\dot{\gamma}|, P)$ is an effective viscosity, which definition is related to the friction coefficient $\mu(I)$ (equation (2)). An important property of the proposed constitutive law is that the effective viscosity diverges to infinity when the shear rate goes to zero. This divergence ensures that a yield criterion exists. Looking at equation (3) in the limit of $|\dot{\gamma}|$ going to zero, we can show that the material flows only if the following condition is satisfied:

$$|\tau| > \mu_s P \text{ where } |\tau| = (0.5\tau_{ij}\tau_{ij})^{0.5} \quad (4)$$

The yield criterion then takes the form of a Drucker–Prager-like criterion²⁴. Below the threshold, the medium behaves locally as a rigid body. It is interesting to note that within this framework, the granular media can be viewed as a visco-plastic fluid²⁵. The specificity compared to classical Bingham or Herschel–Bulkley fluids is that the effective viscosity depends both on the shear rate and on the local pressure. This property is linked to the frictional nature of stresses in granular media.

To test this rheology we performed experiments of granular flows on a heap as sketched in Fig. 2. This set-up is similar to our previous study¹⁹ except that here sidewalls are made rough by gluing one layer of beads on them. This imposes a well-defined no-slip boundary condition at the walls. This configuration represents a severe test for the model, since it gathers in a single configuration several specificities of granular flows. First, when grains are released from the hopper, a steady regime is reached with a strongly sheared layer flowing on top of a static zone. The slope and the thickness of the flowing layer are selected by the system. Second, owing to the rough sidewalls used here, a significant shear exists also in the transverse direction, the flow pattern being thus fully three-dimensional. The experiments are carried out using glass beads 0.53 mm in diameter

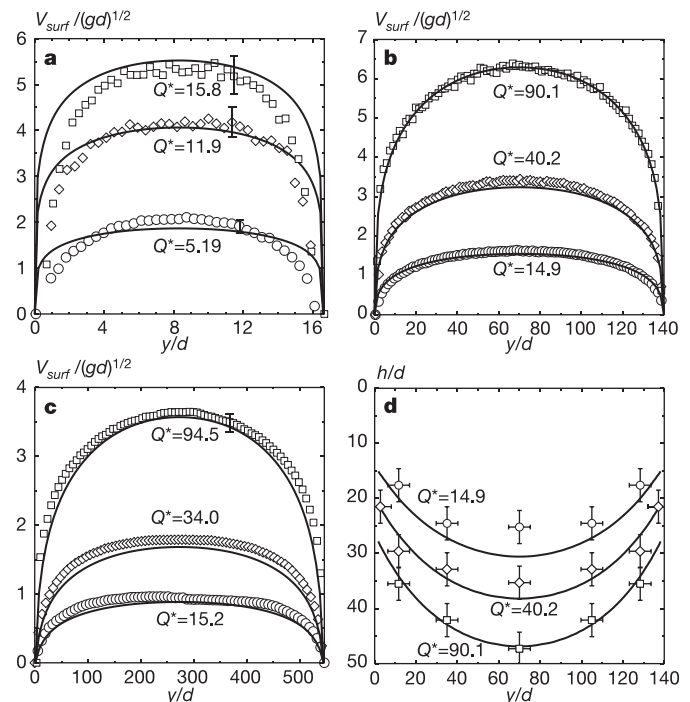


Figure 4 | Comparison of 3D simulations (lines) and experimental results (symbols) for different flow rates ($Q^* = Q/d^{3/2}g^{1/2}$). **a, b, c**, Free-surface velocity profiles for channel width $W = 16.5d$ (**a**), $W = 140d$ (**b**) and $W = 546d$ (**c**). **d**, Depths of the flowing layer across the channel for $W = 140d$. The experimental and computational flow rates are equal within 2.5%. The error bars represent the dispersion of the measurements for different experiments.

and the two control parameters are the width W of the channel and the flow rate per unit of width Q . The present study focuses on the steady and uniform regime characterized by a constant slope and a velocity aligned along the x direction and invariant along the flow (a tiny y component can be observed close to the wall, which remains 20 times smaller than the stream-wise velocity). We performed systematic measurements of the free-surface inclination θ , of the free-surface-velocity profile $V_{\text{surf}}(y)$ using particle-imaging velocimetry, and we get estimates of the thickness of the flowing layer $h(y)$ using an erosion method¹⁹.

To compare the experimental results with the predictions of the local rheology, we perform numerical simulations of a granular fluid described by the constitutive law equation (3) and flowing in an inclined U-shaped channel with a no-slip boundary condition at the three walls. The velocity $u(y, z)$ is assumed to be aligned with x and to depend only on y and z . To get the 3D steady velocity profile, we solve the incompressible Navier–Stokes equations with the internal stress being given by equation (3) using a finite difference scheme. For the rheological parameters μ_s , μ_2 and I_0 coming into play in equation (2), we choose the values given by the experimental data of flows on the inclined planes¹⁸ where the same particles were used (see ref. 19 for how to compute these parameters): $\mu_s = \tan(20.9)$, $\mu_2 = \tan(32.76)$ and $I_0 = 0.279$. This choice means that no fitting parameter will exist when we compare results from the simulations to the experimental data. A typical velocity profile obtained by the model is shown in Fig. 3. We first observe that a static zone develops at the base of the channel. The limit of the static zone varies across the channel, the flowing layer being larger in the centre than close to the walls. The second observation is that the velocity profile is truly 3D and sheared in both y and z directions.

We then tried to quantitatively compare the velocity profiles predicted by the simulations with the ones measured experimentally. In the simulation we impose the inclination and compute the flow

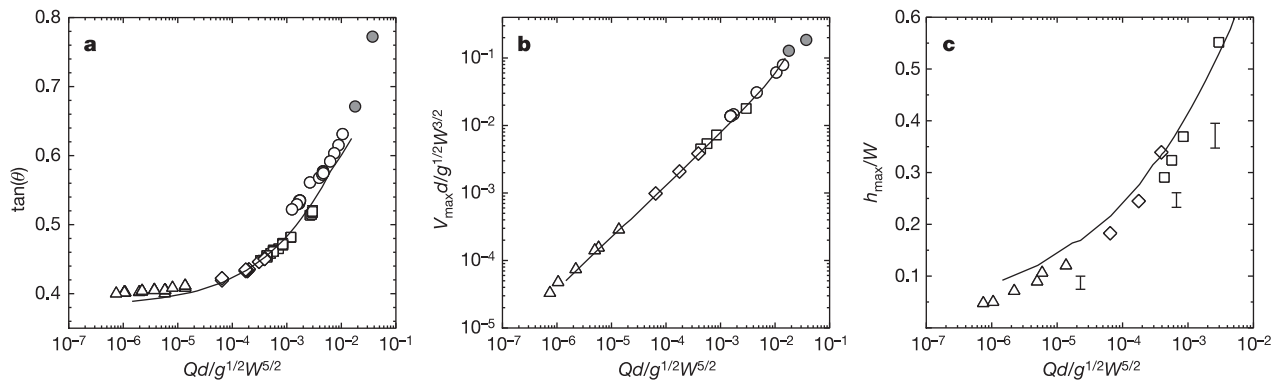


Figure 5 | Scaling laws for the experimental measurements (symbols) compared to the predictions of the model (lines). Free-surface inclination $\tan(\theta)$ (a), rescaled maximum free-surface velocity $V_{\max}$ (b) and maximum flowing thickness $h_{\max}/W$ (c), as a function of the rescaled flow rate Q' . Data collapse onto single curves for all flow rates and all widths: $W = 16.5d$ (circles), $W = 55d$ (squares), $W = 140d$ (diamonds), $W = 546d$ (triangles).

rate *a posteriori*, whereas in the experiments, the flow rate is controlled and the inclination is measured. Figures 4a–c show the free-surface velocity profiles obtained in the experiments and simulations for different widths and different flow rates. The experimental data are the symbols and the continuous lines are the prediction of the 3D rheology. The agreement is good and quantitative. A slight deviation between experiment and model is observed in the narrower channel, 16.5 particle-diameters wide. In Fig. 4d, we also compare the prediction of the theory for the thickness of the flowing layer. In both theory and experiments, the flowing layer is thicker in the centre than at the walls. A quantitative agreement is again observed, although the simulation systematically overestimates the flowing layer thickness. This could be due to the not-very-precise erosion method used for estimating the thickness. All these results show that the proposed rheology gives quantitative predictions for this complex 3D flow, a striking success for a model that has been entirely calibrated based on a different flow configuration.

We have systematically carried out experiments for a wide range of flow rates and channel widths, and within 15% a quantitative agreement is always observed. To compare experiments and simulations in a systematic way, it is interesting to notice that simple scalings can be predicted from the rheology proposed. It is easy to show analytically that one can get rid of the width of the channel in the equations of motion by using the following dimensionless variables: $z' = z/W$, $y' = y/W$, $V' = Vd/g^{0.5}W^{1.5}$ and $Q' = Qd/g^{0.5}W^{2.5}$ (see Supplementary Information). It follows that the inclination of the pile θ , the maximum velocity in the centreline of the channel $V'_{\max}$ and the maximum flowing thickness $h'_{\max}$ should all depend only on Q' . In Fig. 5, we show that the experimental measurements follow the predicted scaling and that the numerical simulations (continuous lines in Fig. 5) give quantitative predictions. One interesting result of this scaling analysis is that the thickness h scales with the width W , meaning that neither the thickness of the flowing layer nor the shear rate are intrinsic properties of the granular media but are controlled by the width of the channel and the flow rate.

We conclude that the simple visco-plastic constitutive law proposed seems to describe dense granular flows very well. Once calibrated on the inclined plane configuration, the model quantitatively captures the complex 3D sheared flow observed when grains flow in between two rough walls. Limits of the approach exist that mainly concern the yield criterion. Within the proposed constitutive law, the flow threshold is simply described by a Coulomb criterion. However, the transition between solid-like and liquid-like behaviour in granular matter seems much more complex, involving shear bands^{26,27}, intermittent flows²⁸ and hysteretic phenomena^{29,30}. Such

In c, the error bars refer to the dispersion of the measurements for each width (the dispersions are smaller than the size of the symbols in a and b). The grey circles (a, b) refer to measurements above the critical angle $\theta_2 = \arctan \mu_2$. In this case, no steady uniform flow is predicted by the model. Experimentally, we observe a transition to a collisional regime (see Supplementary Information and Supplementary Movies).

features should be included in a more comprehensive rheology. However, we believe that the simple visco-plastic rheology presented here represents a minimal model that quantitatively captures the basic features of granular flows important in many applications. We think this model could help to take into account more accurately the complex yield features specific to granular matter.

Received 25 November 2005; accepted 12 April 2006.

1. Savage, S. B. Analysis of slow high-concentration flows of granular materials. *J. Fluid Mech.* **377**, 1–26 (1998).
2. Mills, P., Loggia, D. & Texier, M. Model for stationary dense granular flow along an inclined wall. *Europhys. Lett.* **45**, 733–738 (1999).
3. Aranson, I. S. & Tsimring, L. S. Continuum description of avalanches in granular media. *Phys. Rev. E* **64**, 020301 (2001).
4. Pouliquen, O., Forterre, Y. & Le Dizes, S. Slow dense granular flows as a self-induced process. *Adv. Complex Syst.* **4**, 441–450 (2001).
5. Bocquet, L., Losert, W., Schalk, D., Lubensky, T. C. & Gollub, J. P. Granular shear flow dynamics and forces: Experiment and continuous theory. *Phys. Rev. E* **65**, 011307 (2002).
6. Ertas, D. & Halsey, T. C. Granular gravitational collapse and chute flow. *Europhys. Lett.* **60**, 931–937 (2002).
7. Lemaître, A. Origin of a repose angle: kinetics of rearrangement for granular materials. *Phys. Rev. Lett.* **89**, 064303 (2002).
8. Mohan, L. S., Rao, K. K. & Nott, P. R. A frictional cosserat model for the slow shearing of granular materials. *J. Fluid Mech.* **457**, 377–409 (2002).
9. Josserand, C., Lagrée, P.-Y. & Lhuillier, D. Stationary shear flows of dense granular materials: a tentative continuum modelling. *Eur. Phys. J. E* **14**, 127–135 (2004).
10. Kumaran, V. Constitutive relations and linear stability of a sheared granular flow. *J. Fluid Mech.* **506**, 1–43 (2004).
11. Jaeger, H. M., Nagel, S. R. & Behringer, R. P. Granular solids, liquids, and gases. *Rev. Mod. Phys.* **68**, 1259–1273 (1996).
12. Goldhirsch, I. Rapid granular flows. *Annu. Rev. Fluid Mech.* **35**, 267–293 (2003).
13. Nedderman, R. M. Static and kinematics of granular materials. (Cambridge Univ. Press, Cambridge, UK, 1992).
14. GDR MiDi group, On dense granular flows. *Eur. Phys. J. E* **14**, 341–365 (2004).
15. Da Cruz, F., Emam, S., Prochnow, M., Roux, J.-N. & Chevoir, F. Rheophysics of dense granular materials: Discrete simulation of plane shear flows. *Phys. Rev. E* **72**, 021309 (2005).
16. Iordanoff, I. & Khonsari, M. M. Granular lubrication: toward an understanding between kinetic and fluid regime. *ASME J. Tribol.* **126**, 137–145 (2004).
17. Pouliquen, O. & Forterre, Y. Friction law for dense granular flows: application to the motion of a mass down a rough inclined plane. *J. Fluid Mech.* **453**, 133–151 (2002).
18. Forterre, Y. & Pouliquen, O. Long-surface-wave instability in dense granular flows. *J. Fluid Mech.* **486**, 21–50 (2003).
19. Jop, P., Forterre, Y. & Pouliquen, O. Crucial role of sidewalls in dense granular flows: consequences for the rheology. *J. Fluid Mech.* **541**, 167–192 (2005).
20. Savage, S. B. The mechanics of rapid granular flows. *Adv. Appl. Mech.* **24**, 289–366 (1984).
21. Ancey, C., Coussot, P. & Evesque, P. A theoretical framework for very concentrated granular suspensions in a steady simple shear flow. *J. Rheol.* **43**, 1673–1699 (1999).

22. Pouliquen, O. Scaling laws in granular flows down rough inclined planes. *Phys. Fluids* **11**, 542–548 (1999).
23. Silbert, L. E. *et al.* Granular flow down an inclined plane: Bagnold scaling and rheology. *Phys. Rev. E* **64**, 051302 (2001).
24. Drucker, D. C. & Prager, W. Soil mechanics and plastic analysis of limit design. *Q. Appl. Math.* **10**(2), 157–175 (1952).
25. Tanner, R. I. *Engineering Rheology* (Clarendon Press, Oxford, USA, 1985).
26. Mueth, D. M. *et al.* Signature of granular microstructure in dense shear flows. *Nature* **406**, 385–389 (2000).
27. Fenistein, D. & Van Hecke, M. Wide shear zones in granular bulk flow. *Nature* **425**, 256 (2003).
28. Howell, D., Behringer, R. P. & Veje, C. Stress fluctuations in a 2D granular couette experiment: a continuous transition. *Phys. Rev. Lett.* **82**, 5241–5244 (1999).
29. Daerr, A. & Douady, S. Two types of avalanche behaviour in granular media. *Nature* **399**, 241–243 (1999).
30. Börzsönyi, T., Halsey, T. C. & Ecke, R. E. Two scenarios for avalanche dynamics in inclined granular layers. *Phys. Rev. Lett.* **94**, 208001 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by grants from CEFIPRA and from the French ANR (PIGE project).

Author Contributions P.J. performed all experimental work and numerical simulations; P.J., Y.F. and O.P. analysed results and co-wrote the paper.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.J. (Pierre.Jop@polytech.univ-mrs.fr).

Synthesis and structural analysis of 2-quinuclidonium tetrafluoroborate

Kousuke Tani¹ & Brian M. Stoltz¹

The amide functional group is one of the most fundamental motifs found in chemistry and biology, and it has been studied extensively for the past century¹. Typical acyclic amides are planar. But the amide groups of bicyclic bridgehead lactams^{1,2} are highly twisted, and this distortion from planarity can dramatically affect the stability and reactivity of these amides; it also increases the basicity of the nitrogen so that it often behaves more like an amine than a typical planar amide. As a result, the structures and reactivity profiles of these 'anti-Bredt' amides³ differ significantly from those of planar amides. It is possible that this twisting phenomenon is not exclusive to cyclic systems—non-planarity may also be a critical biological design element that leads to amide ground-state destabilization and alters the reactivity, selectivity and mechanism of various protein and enzymatic processes (such as amide hydrolysis^{4–10}). The intriguing qualities of these twisted amides were first recognized in 1938 (ref. 11), wherein one of the simplest families was introduced—molecules containing the 1-azabicyclo[2.2.2]octan-2-one system. But the parent member of this group, 2-quinuclidone (molecule 1 in this paper), has not yet been unambiguously synthesized. Here, we report the chemical synthesis, isolation and full characterization of the HBF₄ salt of 1. Critical to the success of the synthesis and isolation was the decision to generate 1 by a route other than classical amide bond formation. We anticipate that these results will provide a greater understanding of the properties of amide bonds.

Following the initial description of 2-quinuclidone (molecule 1) by Lukeš¹¹, Woodward became interested in 1 as it related to the structure and synthesis of quinine (in ~1941) and later in the context of the structural assignment of penicillin and the predicted increase in chemical reactivity of both amides due to non-planarity arguments^{12,13}. Woodward connected his laboratory's difficulty in preparing bicyclic lactam 1 from the corresponding amino acid (Fig. 1a) to the unstable and unusual nature of the twisted amide. This was in contrast to the ease of ring closure for the constitutional isomer (Fig. 1b)¹². The marked electrophilicity of penicillin was thus attributed not only to the unusual β-lactam structure, but also to the non-planar distortions of the amide in the fused bicycle.

Then in 1957 Yakhontov claimed to have synthesized 1¹⁴, using the Woodward strategy of amide bond formation. Activation of the amino acid as the acid chloride followed by treatment with mild base and an aqueous work-up was reported to give 1, characterized only by elemental analysis for nitrogen. Subsequently, Pracejus¹⁵ failed to isolate pure 1 by Yakhontov's method, calling the original synthesis into question. The preparation and isolation by Pracejus of a number of methyl-substituted 2-quinuclidone derivatives (that is, molecules 2–5; Fig. 1c) using a similar strategy lend support to the notion that the original Yakhontov synthesis was flawed, at least in the isolation of 1^{16,17}. Indeed, the unambiguous preparation of the parent 2-quinuclidone structure with definitive characterization remains

elusive^{18–24}. Despite this, *ab initio* molecular orbital calculations on 1 and its *N*- and *O*-protonated forms have been reported by Greenberg^{19,25}. Owing to its colourful history, and the apparent challenges related to its preparation, isolation, and characterization, we wished to address the synthesis of 1. Further, given the difficulties associated with the classical route to 1, and in particular the isolation of the intact molecule, we chose to use an alternative strategy.

The core of our design of the synthesis of 1 was to use reactions involving dinitrogen (N₂) release as a driving force for the construction of the strained bicyclic core. This would potentially obviate aqueous work-up conditions in the final step, a manoeuvre that haunted the original report by Yakhontov. Two distinct retrosynthetic pathways were envisioned to construct the 1-azabicyclo[2.2.2]octane skeleton (Fig. 2a). In the first plan (Fig. 2a, route I), a C–C bond formation via C–H insertion of an acylcarbene species could be catalysed by Rh^{II} de-diazotization of the known acyldiazo derivative 6. Promisingly, acylcarbene chemistry has been used for the construction of the C–N bond in other highly twisted bridgehead lactams²⁶. Our second approach involves C–N bond formation using an intramolecular Schmidt–Aubé reaction²⁷ of cyclopentanone azide derivative 7 (Fig. 2a, route II). This intramolecular reaction was first reported by Aubé in 1991 and its use for the synthesis of

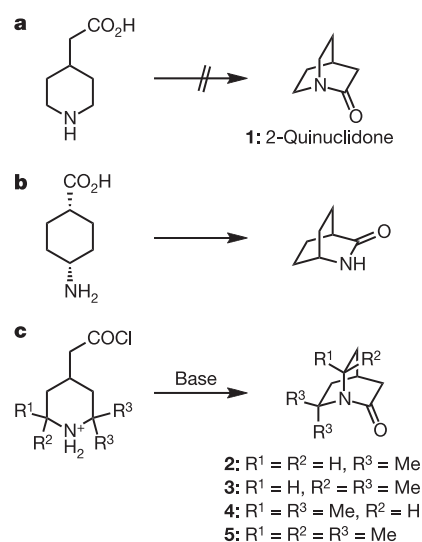


Figure 1 | Ring closure reactions to form bicyclic lactams. a, Intramolecular amide bond formation fails to deliver 2-quinuclidone (1). **b**, Intramolecular amide bond formation readily produces a structural isomer of 1. **c**, Intramolecular lactamization produces a number of substituted variants of 1.

¹The Arnold and Mabel Beckman Laboratories of Chemical Synthesis, Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 E. California Boulevard, M/C 164-30, Pasadena, California 91125, USA.

N-substituted lactams has been established over the past 15 years. Recently, this method has been applied to the construction of another type of twisted lactam²⁸.

Our initial attempts towards the synthesis of 2-quinuclidone (**1**) focused on the C–H insertion of rhodium-stabilized acylcarbenes generated from diazocarbonyl derivatives. These attempts led predominantly to complex mixtures of largely intractable materials. In one attempt, such a reaction (**6**, 2 mol% Rh₂(OAc)₄, CH₂Cl₂, 23 °C) produced small amounts of isomeric olefins **8** (26% yield) resulting from dimerization of the starting material (Fig. 2b). Unfortunately, we did not observe amide **1** or products consistent with its decomposition in any case we examined.

We then turned our attention to the second strategy, namely the use of the intramolecular Schmidt–Aubé reaction. Ketoazide **7** was

prepared as shown in Fig. 2c. Bicyclic lactone **10** was obtained by Baeyer–Villiger oxidation of norcamphor (**9**) and subsequently reduced with LiAlH₄ to give *cis*-diol **11**. Selective tosylation of diol **11** followed by azide substitution of tosylate **12** furnished azidoalcohol **13**. Oxidation of hydroxyazide **13** with Dess–Martin periodinane (DMP) produced the desired ketoazide **7** in 93% yield.

Next, we exposed ketoazide **7** to typical intramolecular Schmidt–Aubé reaction conditions, namely, neat trifluoroacetic acid (TFA) at 60 °C (Fig. 2c). Immediate gas evolution was observed and **7** was consumed completely within 3 h. As the product was expected to be unstable due to the lability of the amide, the resulting solution was treated with methanol. Subsequent Boc-protection and purification afforded derivatives **16** and **17** in 56% and 34% yield, respectively. The production of **16** and **17** strongly suggested initial formation of the desired 2-quinuclidone (**1**), along with an undesired isomer. Additionally, the ¹H NMR spectrum of the crude reaction mixture before methanolysis displayed resonances consistent with a mixture of bicyclic lactams, favouring the quinuclidone ring system.

To improve the selectivity of the rearrangement, we carried out further optimization of the reaction conditions by screening various acids and solvents. The result of each experiment was evaluated after treatment of the crude reaction mixture with MeOH. While neat TFA yielded the desired rearrangement product (Fig. 2c), employing 2 equiv of TFA in CH₂Cl₂ as solvent afforded only starting ketoazide. Alternatively, 2 equiv of trifluoromethanesulphonic acid (TfOH) in CH₂Cl₂ at 20 °C gave rearrangement in almost quantitative yield, although the selectivity did not improve. Other acids, such as Tf₂NH and HBF₄, were also examined in CH₂Cl₂ solvent, but led to similar results. In contrast, the choice of solvent displayed a moderate effect on selectivity. The use of HBF₄ in toluene gave the rearrangement products **14** and **15** in the ratio 68:32, while in Et₂O the ratio was 76:24. Although we attempted further optimization, HBF₄ in Et₂O at 20 °C remained the best conditions for maximizing the production of **14**, and by inference, **1**, as its protonated derivative.

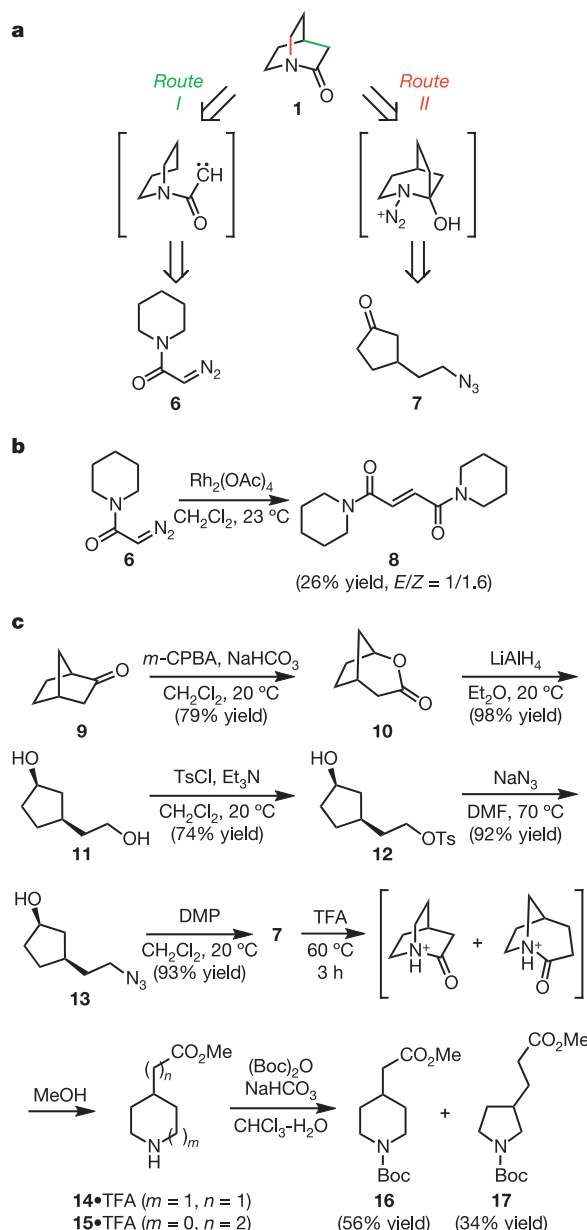


Figure 2 | Strategies for the synthesis of 2-quinuclidone. **a**, Retrosynthetic analysis of 2-quinuclidone (**1**). **b**, Attempted rhodium(II)-catalysed C–H insertion. **c**, Preparation of ketoazide **7** and use in the intramolecular Schmidt–Aubé reaction. *m*-CPBA, *meta*-chloroperoxybenzoic acid; TsCl, *para*-toluenesulfonyl chloride; DMF, *N,N*-dimethyl formamide; DMP, Dess–Martin periodinane; TFA, trifluoroacetic acid; Boc, *tert*-butoxycarbonyl.

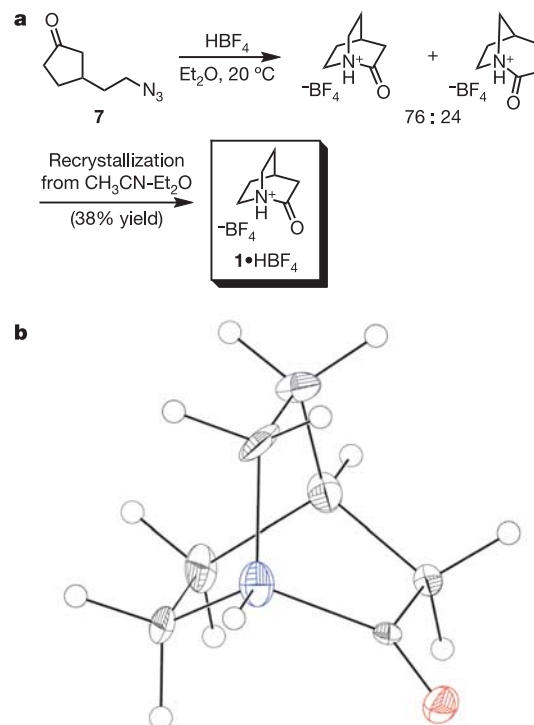


Figure 3 | Isolation and unambiguous characterization of 2-quinuclidonium tetrafluoroborate. **a**, Isolation of **1**•HBF₄. **b**, ORTEP drawing of **1**•HBF₄ (shown with 50% probability ellipsoids, BF₄ omitted for clarity). Nitrogen = blue, oxygen = red, carbon = black and hydrogen = white).

Table 1 | Comparison of structural parameters for **1 and formamide**

Compound	Bond length (Å)		Distortion parameters ²⁹		
	N–C(O)	C=O	χ_N	χ_C	τ
1 ·HBF ₄ (X-ray); this work	1.526(5)	1.192(4)	59.5°	0.2°	90.9°
	1.484(6)	1.168(6)	58.9°	–2.4°	90.8°
1 (N-protonated, calculated) ^{19,25}	1.504	1.167	57.6°	0.0°	89.9°
1 (calculated) ^{19,25}	1.433	1.183	55.6°	0.0°	90.0°
Formamide (planar, calculated) ^{19,25,30}	1.349	1.193	0.0°	0.0°	0.0°
Formamide (perpendicular, calculated) ^{19,25,30}	1.423	1.179	63.4°	0.0°	90.0°

In this work, two crystallographically independent molecules were observed in the unit cell.

Because of the modest selectivity of the rearrangement, we focused on direct purification of the ~3:1 crude mixture of bicyclic lactams. A key observation was that in the reaction of ketoazide **7** with HBF₄ in Et₂O, following rearrangement, a precipitate formed. Gratifyingly, recrystallization of the crude solid from CH₃CN–Et₂O produced pure 2-quinuclidonium tetrafluoroborate (**1**·HBF₄) as colourless crystals in 38% overall yield from ketoazide **7** (Fig. 3a).

The structure of **1**·HBF₄ was unambiguously determined by spectroscopic evaluation. The carbonyl infrared absorption band of **1**·HBF₄ was observed at 1,822 cm^{–1} (KBr). This value was comparable to those of the hydrochloride salts of **3**¹⁵ and **5**¹⁶ (1,818 cm^{–1} and 1,811 cm^{–1}, respectively). Interestingly, this value is much closer to that of an acid chloride or anhydride derivative than to an amide or an aminoketone. Meanwhile, the ¹³C NMR chemical shift of the carbonyl group was observed at 175.9 p.p.m. (CD₃CN). Crystals suitable for X-ray analysis were obtained by further recrystallization from CH₃CN–Et₂O. Figure 3b shows the X-ray crystallographic structure of **1**·HBF₄. All of the hydrogen atoms were apparent from the electron density map. The structure shows the N-protonated form of **1** and pyramidalization at nitrogen. This is in good agreement with Greenberg's calculated proton affinity (PA) of **1**, which was found to favour the N-protonated form by 24.1 kcal mol^{–1} relative to the O-protonated form²⁵.

Table 1 summarizes selected bond lengths and distortion parameters from X-ray crystallography and the previously calculated results. Two crystallographically independent molecules corresponding to **1**·HBF₄ were observed in the unit cell. The observed lengths of the N–C(O) bonds in **1**·HBF₄ were 1.484 and 1.526 Å while the C=O bond distances were 1.168 and 1.192 Å. These distances were in good agreement with the calculated results of N-protonated **1**^{19,25}. Dunitz and Winkler reported three independent distortion parameters, χ_N , χ_C , and τ , for quantitative evaluation of the twisting of an amide bond²⁹. These parameters describe the pyramidalization at the nitrogen (χ_N) and carbon atoms (χ_C) and the torsion about the C–N bond (τ), respectively. For the representative planar amide, formamide, these parameters are all 0° (ref. 30). In contrast, perpendicular formamide was calculated to have a χ_N of 63.4° and τ of 90.0°. For the two independent structures of **1**·HBF₄ found in the unit cell, χ_N was calculated to be 58.9° and 59.5°, while τ was found to be 90.8° and 90.9°. These results are in reasonable accord with the perpendicular formamide example and quantitatively illustrate the highly twisted nature of **1**·HBF₄.

Recently, we have started to investigate the chemical reactivity of 2-quinuclidonium tetrafluoroborate. For instance, **1**·HBF₄ reacts instantaneously in neat deuterated water, D₂O. In the ¹H NMR spectrum acquired 1 min after dissolving **1**·HBF₄ in D₂O, only the ring-opened amino acid was observed, corresponding to a $t_{1/2} < 15$ s. The reaction with 5 equiv of D₂O in CD₃CN (0.084 M) was slower, with $t_{1/2} = 135$ min. Additionally, **1**·HBF₄ is unstable to manipulations in common nucleophilic solvents (for example, DMSO, pyridine, MeOH) aside from CH₃CN and Et₂O. Attempts to observe **1** by free-basing **1**·HBF₄ have been unsuccessful, and have resulted largely in polymeric material (see Supplementary Information for details). These observations illustrate the propensity for this unusual amide linkage to undergo nucleophile-induced cleavage and highlight the fragile nature of this interesting molecule.

In conclusion, we have achieved the first unambiguous synthesis, isolation, and X-ray crystallographic characterization of 2-quinuclidone as its tetrafluoroborate salt (**1**·HBF₄). Our synthesis provides access to the parent compound of the quinuclidone series nearly 70 years after its conception as an 'anti-Bredt' amide by Luke¹¹. Critical to the success of the synthesis and isolation was the decision to generate **1** by a route other than classical amide bond formation. Our ability to prepare and characterize 2-quinuclidonium tetrafluoroborate illustrates the power of the Schmidt–Aubé reaction for the synthesis of highly strained amides. Additionally, these conditions furnish the N-protonated amide directly, saving the molecule from self-destruction. We expect that the results presented here will provide a greater understanding of the properties of amide bonds.

METHODS

Routine synthetic experimental techniques were used for the preparation, isolation, and analysis of all new compounds as described in the Supplementary Information.

Procedure for the synthesis of 2-quinuclidonium tetrafluoroborate. See Fig. 3a for the isolation reaction of **1**·HBF₄. A 10 ml tube equipped with stirbar and three-way stopcock was flame-dried under vacuum, backfilled with dry nitrogen, and charged with ketoazide **7** (306 mg, 2.00 mmol, 1.0 equiv) and dry ether (4 ml). To this solution we added ethereal HBF₄ (54 wt%, 0.55 ml, 4.00 mmol, 2.0 equiv) at 0 °C and the resulting mixture was stirred at ambient temperature (20 °C) until gas evolution ceased (3 h). The supernatant of the resulting suspension was removed by syringe and the remaining white solid was washed with dry ether (3 × 3 ml) and dried under vacuum. The resulting crude solid was then dissolved in 4 ml of dry acetonitrile and transferred to a 10 ml test tube, which was placed in a septum-sealed 200 ml Erlenmeyer flask. Dry Et₂O (10 ml) was then added to the Erlenmeyer flask outside the tube, and the resulting flask was settled in a desiccator (P₂O₅) at ambient temperature until crystals formed (6 days). After the mother liquor was removed by syringe, the solid was washed with dry Et₂O (3 × 5 mL) and dried under vacuum to yield 164 mg (0.770 mmol, 38% yield) of 2-quinuclidonium tetrafluoroborate **1**·HBF₄ as colourless crystals.

Recrystallization from CH₃CN–Et₂O provided suitable crystals for X-ray analysis; m.p.: 185–200 °C decomposition; ¹H NMR (300 MHz, CD₃CN) δ 8.02 p.p.m. (br, 1H), 3.85–3.60 (m, 4H), 2.99 (d, $J = 3.0$ Hz, 2H), 2.51 (m, 1H), 2.10–1.90 (m, 4H); ¹³C NMR (75 MHz, CD₃CN) δ 175.9 p.p.m., 48.1, 40.1, 25.7, 22.7; IR (KBr) 3,168 cm^{–1}, 2,981 cm^{–1}, 1,822 cm^{–1}, 1,468 cm^{–1}, 1,398 cm^{–1}; HRMS (FAB) (m/z): calculated for C₇H₁₂NO [M + H]⁺ 126.0919, found 126.0920.

Received 21 March; accepted 26 April 2006.

- Greenberg, A., Breneman, C. M. & Liebman, J. F. *The Amide Linkage: Selected Structural Aspects in Chemistry, Biochemistry, and Materials Science* (Wiley, New York, 2000).
- Greenberg, A. Twisted bridgehead bicyclic lactams. *Mol. Struct. Energetics* **7**, 139–178 (1988).
- Bredt, J. Über sterische Hinderung in Brückenringen (Bredtsche regel) und über die meso-trans-Stellung in kondensierten Ringsystemen des Hexamethylens. *Liebigs Ann. Chem.* **437**, 1–13 (1924).
- Mujika, J. I., Mercero, J. M. & Lopez, X. Water-promoted hydrolysis of a highly twisted amide: rate acceleration caused by the twist of the amide bond. *J. Am. Chem. Soc.* **127**, 4445–4453 (2005).
- Romanelli, A., Shekhtman, A., Cowburn, D. & Muir, T. W. Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. *Proc. Natl Acad. Sci. USA* **101**, 6397–6402 (2004).
- Poland, B. W., Xu, M. Q. & Quijcho, F. A. Structural insights into the protein splicing mechanism of PI-SceI. *J. Biol. Chem.* **275**, 16408–16413 (2000).

7. Brown, R. S., Bennet, A. J. & Slebocka-Tilk, H. Recent perspectives concerning the mechanism of H_3O^+ - and hydroxide-promoted amide hydrolysis. *Acc. Chem. Res.* **25**, 481–488 (1992).
8. Liu, J., Albers, M. W., Chen, C. M., Schreiber, S. L. & Walsh, C. T. Cloning, expression, and purification of human cyclophilin in *Escherichia coli* and assessment of the catalytic role of cysteines by site-directed mutagenesis. *Proc. Natl Acad. Sci. USA* **87**, 2304–2308 (1990).
9. Harrison, R. K. & Stein, R. L. Mechanistic studies of peptidyl prolyl cis-trans isomerase: evidence for catalysis by distortion. *Biochemistry* **29**, 1684–1689 (1990).
10. Mock, W. L. Torsional-strain considerations in enzymology: Some applications to proteases and ensuing mechanistic consequences. *Bioorg. Chem.* **5**, 403–414 (1976).
11. Lukeš, R. Sur une nouvelle application de la règle de Bredt. *Coll. Czech. Chem. Commun.* **10**, 148–152 (1938).
12. Johnson, J. R., Woodward, R. B. & Robinson, R. The constitution of the penicillins. In *The Chemistry of Penicillin* (eds Clark, H. T., Johnson, J. R. & Robinson, R.) 440–454 (Princeton Univ. Press, Princeton, 1949).
13. Wasserman, H. H. Profile and scientific contributions of Professor R. B. Woodward. *Heterocycles* **7**, 1–15 (1977).
14. Yakhontov, L. N. & Rubsitov, M. V. The synthesis of quinuclidone-2. *J. Gen. Chem. USSR* **27**, 83–87 (1957).
15. Pracejus, H., Kehlen, M., Kehlen, H. & Matschiner, H. Neues zur sterischen Mesomeriehemmung bei Lactamen vom typ des α -Chinuclidons. *Tetrahedron* **21**, 2257–2270 (1965).
16. Levkoeva, E. I., Nikitskaya, E. S. & Yakhontov, L. N. Reactions of 2-quinuclidones with and without scission of C–N bonds. *Dokl. Akad. Nauk SSSR* **192**, 342–345 (1970).
17. Greenberg, A., Wu, G., Tsai, J.-C. & Chiu, Y.-Y. Improved synthesis of 6,6,7,7-tetramethyl-1-azabicyclo[2.2.2]octan-2-one and its stability toward base-induced methanolysis. *Struct. Chem.* **4**, 127–129 (1993).
18. Boivin, J., Gaudin, D., Labrecque, D. & Jankowski, K. Gif oxidation of some alicyclic amines. *Tetrahedr. Lett.* **31**, 2281–2282 (1990).
19. Greenberg, A., Moore, D. T. & DuBois, T. D. Small and medium-sized bridgehead bicyclic lactams: a systematic ab initio molecular orbital study. *J. Am. Chem. Soc.* **118**, 8658–8668 (1996).
20. Blackburn, G. M., Skaife, C. J. & Kay, I. T. Strain effects in acyl transfer reactions. Part 5. The kinetics of hydrolysis of benzoquinuclidin-2-one: a torsionally-distorted amide. *J. Chem. Res. (Miniprint)* 3650–3669 (1980).
21. Bennet, A. J. *et al.* Relationship between amidic distortion and ease of hydrolysis in base. If amidic resonance does not exist, then what accounts for the accelerated hydrolysis of distorted amides? *J. Am. Chem. Soc.* **112**, 6383–6385 (1990).
22. Wang, Q. P., Bennet, A. J., Brown, R. S. & Santarsiero, B. D. Distorted amides as models for activated peptide N–C=O units. 2. The synthesis, hydrolytic profile, and molecular structure of 3,4-dihydro-2-oxo-1,4-propanoquinoline. *Can. J. Chem.* **68**, 1732–1739 (1990).
23. Buchanan, G. L. *et al.* Conformational study of the 1-azabicyclo[3.3.1]nonan-2-one system. Molecular-mechanics calculations and X-ray structure of 5-phenyl-1-azabicyclo[3.3.1]nonan-2-one. *J. Chem. Soc. Perkin Trans. 2*, 1709–1712 (1983).
24. Kirby, A. J., Komarov, I. V., Wothers, P. D. & Feeder, N. The most twisted amide: structure and reactions. *Angew. Chem. Int. Edn Engl.* **37**, 785–786 (1998).
25. Greenberg, A. & Venanzi, C. A. Structures and energetics of two bridgehead lactams and their N- and O-protonated forms: an ab initio molecular orbital study. *J. Am. Chem. Soc.* **115**, 6951–6957 (1993).
26. Williams, R. M. & Lee, B. H. Synthesis of a 1,3-bridged β -lactam: a novel, anti-Bredt β -lactam. *J. Am. Chem. Soc.* **108**, 6431–6433 (1986).
27. Aubé, J. & Milligan, G. L. Intramolecular Schmidt reaction of alkyl azides. *J. Am. Chem. Soc.* **113**, 8965–8966 (1991).
28. Lei, Y., Wroblewski, A. D., Golden, J. E., Powell, D. R. & Aubé, J. Facile C–N cleavage in a series of bridged lactams. *J. Am. Chem. Soc.* **127**, 4552–4553 (2005).
29. Dunitz, J. D. & Winkler, F. K. Amide group deformation in medium-ring lactams. *Acta Crystallogr.* **B31**, 251–263 (1975).
30. Wiberg, K. B. & Laidig, K. E. Barriers to rotation adjacent to double bonds. 3. The C–O barrier in formic acid, methyl formate, acetic acid, and methyl acetate. The origin of ester and amide “resonance”. *J. Am. Chem. Soc.* **109**, 5935–5943 (1987).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank H. H. Wasserman, E. J. Corey, and N. J. Leonard for discussions. We thank Ono Pharmaceutical Co. Ltd (for a postdoctoral fellowship to K.T.), the Research Corporation, the Camille and Henry Dreyfus Foundation, Merck, Pfizer, Lilly, Amgen and Bristol-Myers Squibb for financial support. We also thank M. W. Day and L. M. Henling for their X-ray crystallographic expertise.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. Crystallographic data have been deposited at the CCDC (12 Union Road, Cambridge CB2 1EZ, UK) and copies can be obtained on request, free of charge, by quoting the publication citation and the deposition number 296767. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to B.M.S. (stoltz@caltech.edu).

Decay of aftershock density with distance indicates triggering by dynamic stress

K. R. Felzer¹ & E. E. Brodsky²

The majority of earthquakes are aftershocks¹, yet aftershock physics is not well understood. Many studies suggest that static stress changes^{2,3} trigger aftershocks, but recent work suggests that shaking (dynamic stresses) may also play a role^{4,5}. Here we measure the decay of aftershocks as a function of distance from magnitude 2–6 mainshocks in order to clarify the aftershock triggering process. We find that for short times after the mainshock, when low background seismicity rates allow for good aftershock detection, the decay is well fitted by a single inverse power law over distances of 0.2–50 km. The consistency of the trend indicates that the same triggering mechanism is working over the entire range. As static stress changes at the more distant aftershocks are negligible, this suggests that dynamic stresses may be triggering all of these aftershocks. We infer that the observed aftershock density is consistent with the probability of triggering aftershocks being nearly proportional to seismic wave amplitude. The data are not fitted well by models that combine static stress change with the evolution of frictionally locked faults³.

Previous studies of how aftershock density decays with distance from the mainshock have found a range of functions, including power laws and combinations of power laws with constants and exponentials^{6–8}. The ability to study the decay with improved clarity has recently been provided by the publication of large catalogues with precise earthquake locations. We use the Shearer *et al.* relocated 1984–2002 Southern California catalogue⁹. Standard location error is of the order of tens of metres, similar to that obtained in more localized studies using similar relocation techniques^{10,11}. For most of the analysis we use magnitude $M = 2$ –4 mainshocks and $M \geq 2$ aftershocks. We prefer small mainshocks because their large number allows for good statistical averaging and for the use of a small difference between mainshock and aftershock magnitude, which improves catalogue completeness¹². For near field measurements, where larger mainshocks are necessary for appropriate range and precision, we use $M = 5$ –6 mainshocks and $M \geq 3$ aftershocks. Earthquakes are used as mainshocks if they are sufficiently isolated from larger earthquakes in time and space (see Methods). To improve statistics, we combine the aftershocks of each unit magnitude range of mainshocks (Fig. 1).

The spatial density of a point process can be measured in any dimension. For instance, a density could be the number of points per length, per area, or per volume. We choose to measure linear density, that is, the number of aftershocks per length. This is done by collapsing all of the aftershocks onto a single line with their position on the line equal to their distance from the mainshock. The number of aftershocks per unit length can then be measured at different positions using standard statistical tools (see Methods).

We first study earthquakes occurring within 5 min and 50 km of $M = 2$ –4 mainshocks. The short time window minimizes the amount of background seismicity¹³, that is, earthquakes not aftershocks of the

designated mainshocks. We approximate the mainshocks as point sources and measure the distance, r , between mainshock and aftershock hypocentres. From 0.2 to 50 km, the data are well fitted by

$$\rho(r) = cr^{-n} \quad (1)$$

where $n = 1.37 \pm 0.1$ for the $3 \leq M < 4$ mainshocks and 1.35 ± 0.12 for the $2 \leq M < 3$ mainshocks, and c is a constant that varies with the number of aftershocks (Fig. 2). The error is the 98% confidence interval based on 5,000 bootstraps. For $r < 0.2$ km the point source approximation is no longer accurate (Supplementary Figs 1 and 2).

We also check the applicability of equation (1) to longer times. For 30 minutes of post-mainshock data the time window is still short, and an inverse power law, with $n = 1.36 \pm 0.07$, fits the $3 \leq M < 4$ mainshocks. For $2 \leq M < 3$ mainshocks background seismicity begins to level the decay around 16 km, but a clear inverse power law with $n = 1.37 \pm 0.15$ is evident at shorter distances (Supplementary Fig. 3). Background interference is worse for smaller mainshocks because of lower aftershock productivity per mainshock. For 30 minutes to 25 days of post-mainshock earthquakes there is universal deviation from a pure power law, but we find that a combined power law/background function fits the data well, at 95% and 65% confidence (Supplementary Fig. 4). This suggests that the aftershock decay with distance does not change with time.

To study the density of aftershocks within one fault length of the

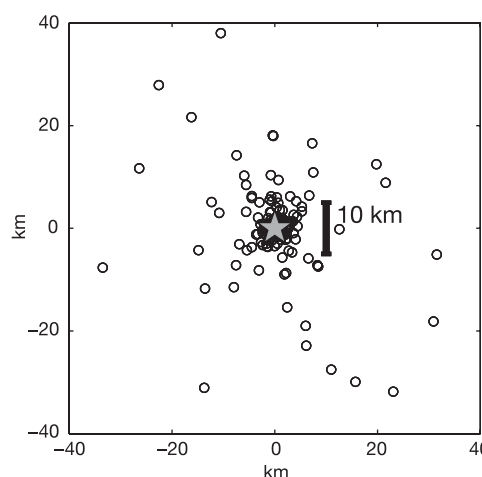


Figure 1 | Combined aftershocks of $M = 3$ –4 mainshocks. To create a composite aftershock data set, we move all of the mainshocks to the origin in space and time and move their aftershocks with them. Data here are for the first 30 minutes of aftershocks of $M = 3$ –4 mainshocks. The grey star gives the locations of the mainshocks, at the origin.

¹US Geological Survey, 525 S. Wilson, Pasadena, California 91106, USA. ²Department of Earth Science, University of California Santa Cruz, 1156 High Street, Santa Cruz, California 95060, USA.

mainshock, we use $M = 5$ – 6 mainshocks that have a Harvard CMT (centroid moment tensor) focal mechanism solution and aftershock distributions that clearly delineate the preferred fault plane. We estimate the length and width of the faults from empirical relationships¹⁴, and centre the fault planes at the median aftershock location. We measure aftershock density from 0.2 km to 12 km from the fault plane (~ 0.05 fault lengths of an $M = 5$ earthquake to one fault length of an $M = 6$ earthquake). As before, we recover an inverse power law, with $n = 1.34 \pm 0.25$ (Fig. 3). The decay levels out at $r < 0.2$ km owing to error in fault plane location and catalogue incompleteness in the very near field.

To verify that the decay we observe is due to aftershock physics, we repeat the analysis for the $M = 2$ – 4 mainshocks with a time-randomized catalogue. This produces a large scatter of points (Fig. 4), indicating that the pattern in Fig. 2 is aftershock-related. To verify that a pure inverse power law is a good functional fit to the aftershocks, we use the Kolmogorov–Smirnov test (see Methods). We also test the fit of a composite power law of the form $\rho(r) = a_1 r^{-n_1} + a_2 r^{-n_2}$. The Bayesian information criterion^{15,16} prefers the single power law fit. To check if our results are catalogue dependent, we verify that aftershocks in unrelocated Japanese and Northern California catalogues also follow an inverse power law decay (Supplementary Fig. 6).

The consistent aftershock decay relationship observed from distances of 0.2 km to 50 km—from within 0.05 fault lengths of $M = 5$ mainshocks to over 100 fault lengths of $M = 2$ – 3 mainshocks—implies that static stress change is not triggering the aftershocks. Static stresses decay rapidly. At 3 km from a $M = 4$ earthquake the static stress change is at most 4 kPa, comparable to tidal stresses that have not been found to trigger earthquakes¹⁷; at 10 km from a $M = 3$ earthquake the static stress change is three orders of magnitude lower. Triggering by static stress in the near field and dynamic stress in the far field would require a discontinuity in the aftershock decay. Only uniform triggering by dynamic stress matches the observation of a single, consistent decay that traverses a wide range of distances.

The hypothesis of aftershock triggering by static stress change has received strong support in part because there is a physical model, rate and state friction³, which explains how static stress changes could result in the power law distribution of aftershock times¹⁸. Our

observations indicate, however, that this model does not fit the distribution of aftershocks in space. At very short times, using a point source approximation for the static stress in a whole space, the model predicts that the density of aftershocks from an earthquake of moment M_0 is

$$\rho(r) = B(r)e^{c_2 r^{-3}} \quad (2)$$

where $B(r)$ is the background seismicity per kilometre per unit time as a function of distance from the mainshock and $c_2 = M_0/4\pi A\sigma$, where σ is normal stress and A is a frictional constant. At a given distance from a set of mainshocks, the observed density in a combined data set is the sum of the individual ones.

The functional form of $B(r)$ can be estimated from distance measurements between random earthquakes (Fig. 4). The scatter is large, but the data can be fitted with the same functional form used by others^{19,20}:

$$B(r) = c_1 r^{(D-1)} \quad (3)$$

This function describes points randomly scattered on a structure with effective dimension D . We substitute equation (3) into equation (2) and use a grid search to find the best fit to the 30 min/16 km aftershocks of $M = 3$ – 4 mainshocks over a wide parameter range (see Methods). For the best fit case, with $D = 0.1$, the summed squared error is 1.4 times worse than for the best inverse power law fit, and the correlation of the data residuals is high ($r = 0.45$; 601 data points) (Supplementary Fig. 7). This correlation has less than a 0.01% chance of occurring for a good functional fit. For the residuals of the inverse power law, $r = 0.0079$. For more realistic values of $D = 1$ – 2 , the rate and state model fit is much worse (Supplementary Fig. 7). The point source approximation may produce an inaccurate representation of the static stress change within one fault length of the source (for example, within 1 km for $M = 4$ mainshocks and 0.1 km for $M = 2$ mainshocks), but the functional shape and associated misfit is problematic at further distances as well.

The poor fit of the data to equation (2) indicates that the aftershocks are not triggered by static stress change coupled with rate and state friction, at least at distances beyond one fault length. We also find more model-dependent evidence that the number of aftershocks triggered varies linearly with dynamic stress change amplitude.

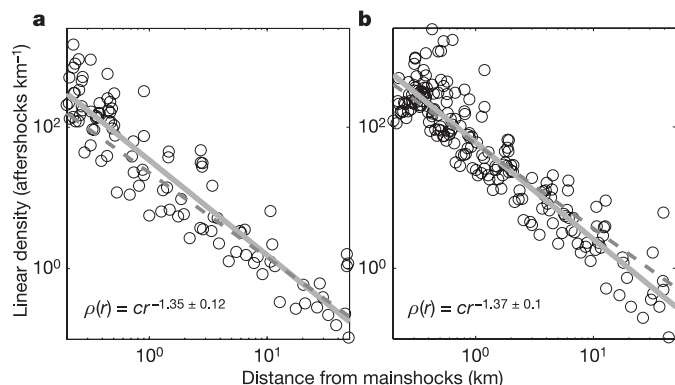


Figure 2 | Distance from the mainshock hypocentre versus aftershock linear density. Aftershocks are $M > 2$ and occur in the first 5 minutes. **a**, $2 \leq M < 3$ mainshocks. The plot uses 7,396 mainshocks and 104 aftershocks. **b**, $3 \leq M < 4$ mainshocks. The plot uses 2,355 mainshocks and 199 aftershocks. The data are fitted with an inverse power law with an exponent of -1.35 ± 0.12 for the $2 \leq M < 3$ data and -1.37 ± 0.1 for the $3 \leq M < 4$ data (solid lines). The fit is made from 0.2 to 50 km for both plots. For comparison, dashed lines give the decay of maximum seismic wave amplitude, a proxy for dynamic stress, as derived from the standard Richter relationship²². Data over a wider distance range are given in Supplementary Fig. 1.

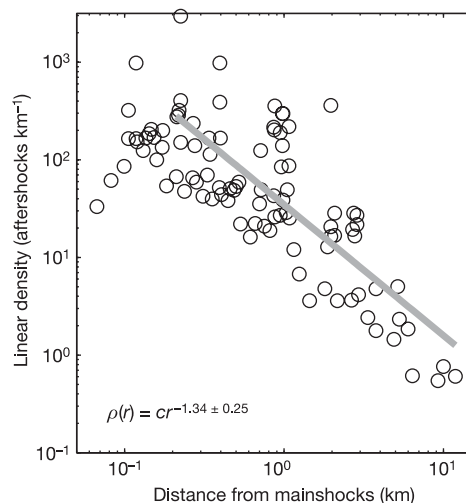


Figure 3 | Aftershock density versus distance from the closest point on the fault planes of $M = 5$ – 6 mainshocks. The plot uses nine mainshocks and 104 $M > 3$ aftershocks that occurred within 2 days of the mainshock. A fit to the data is made from 0.2 to 12 km, or from 0.05 fault lengths of an $M = 5$ earthquake to one fault length of an $M = 6$ earthquake. The decay rate shows good agreement with the far field decay found for $M = 2$ – 4 mainshocks. Far field aftershocks of these mainshocks are given in Supplementary Fig. 5.

Linear aftershock density, $\rho(r)$, can be separated into geometric and physical terms

$$\rho(r) = \frac{N_{\text{aft}}}{\Delta r} = \frac{N_{\text{hyp}}}{\Delta r} \times \frac{N_{\text{aft}}}{N_{\text{hyp}}} = cr^{-1.4} \quad (4)$$

where N_{aft} is the number of aftershocks in a shell of width Δr centred at distance r from the mainshock and N_{hyp} is the number of potential hypocentres, or places where aftershocks could be triggered, in the same shell. If the background seismicity is roughly evenly distributed over the active fault population, the functional form of the background earthquake distribution (equation (3)) can be substituted for the geometric term:

$$\rho(r) = c_1 r^{D-1} c_3 r^m = cr^{-1.4} \quad (5)$$

The trade-off between D and m prevents a direct solution for D . We instead estimate D by fitting the data in Fig. 4. We find a better fit with $D = 1$ than with $D = 2$ or 3 ; that is, the linear density is independent of distance. A D value close to 1 is also suggested by geometric considerations (see Methods).

If $D \approx 1$ the fraction of potential hypocentres triggered decays as $r^{-1.4}$, that is, at a rate slightly stronger than $1/r$. The maximum amplitude of seismic waves²¹ also decay somewhat faster than $1/r$. The standard Richter relationship for maximum short period seismic wave amplitude^{22,23} is well fitted by a combination of $r^{-1.2}$ decay and attenuation with $Q = 300$. Attenuation is less important at the depth of the aftershocks, so the peak amplitude of shaking at depth falls off as $r^{-1.2}$ (Fig. 2). The power law decay of local seismic amplitude is primarily due to the geometrical spreading of the S waves, wave focusing, and surface wave amplitude decay. Given the uncertainty in D , the data are consistent with the probability of triggering an aftershock being proportional to the amplitude of the seismic wave. The probability of aftershock triggering scaling with the amplitude of the mainshock wave also predicts that the number of aftershocks increases by a factor of ten with each mainshock magnitude unit, as has been observed^{20,24,25}.

In summary, the decay of aftershock linear density with distance from $M = 2$ –6 mainshocks is well fitted by an inverse power law. The

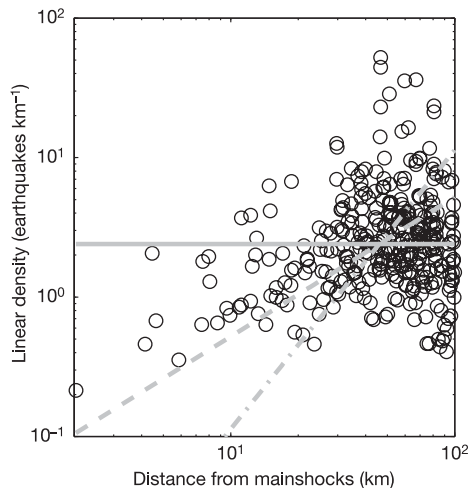


Figure 4 | Distance versus earthquake linear density for a time-randomized catalogue. Distances are measured between ten random earthquakes from the mainshock data set and random catalogue earthquakes, producing 346 earthquake pairs. Unlike the aftershock data, there is no systematic decay of density with distance. The sum of squared residuals is lowest when the data are fitted with the relationship that linear earthquake density is constant (solid line), as opposed to linear density equals r (dashed line), which would correspond to hypocentres being located randomly in two dimensions, or r^2 (dashed-dotted line), which would correspond to random locations in a volume (see equation (3)). The fits are done from 0 to 100 km, which completely covers the range of aftershock data analysed in this paper.

trend exists at least from 0.2 to 50 km for the first 5 minutes of the aftershock sequence. At longer times background seismicity makes distant aftershocks more difficult to detect, but we can trace the decay trend to at least 16 km for the first 30 minutes, and to at least 10 km for the first 2 days. If the linear density of faults is independent of distance ($D \approx 1$), then the data indicate that the probability of triggering an aftershock is directly proportional to the amplitude of seismic shaking. Whether or not $D = 1$, dynamic triggering is preferred by the data. The similarity of aftershock decay from distances of 0.05 to over 100 fault lengths implies a single physical triggering mechanism, and dynamic stress change is the only plausible agent over most of this range.

METHODS

Mainshock and aftershock selection. Earthquakes are used as mainshocks if they are separated from larger earthquakes by at least L km or by t_1 days if the larger earthquake comes first, and t_2 days if it comes after. This separation minimizes contamination from aftershocks of larger earthquakes. L is set at 100 km; larger values (at least up to 500 km) produce the same results. The results are insensitive to the values of t_1 and t_2 as long as $t_1 \ll t < t_2$, where t is the time after the mainshock for which we use aftershock data. For $t = 5$ min and $t = 30$ min we obtain the same aftershock decay for values of t_1 between 3 and 100 days. We use $t_1 = 3$ days, which, within this range, maximizes the number of qualified mainshocks, and $t_2 = 0.5$ days. For Supplementary Fig. 4, where t can be as large as 25 days, we set $t_1 = 100$ days and $t_2 = 26$ days. For Fig. 3, where $t = 2$ days, we use $t_1 = 30$ days and $t_2 = 2$ days.

We use a uniform distance cut-off for measuring aftershocks of all mainshocks, even though the mainshock magnitudes vary. This is justified by our observation, in agreement with previous authors^{8,26}, that the distribution of aftershock distances is independent of mainshock magnitude (Supplementary Fig. 8).

Linear density measurement. To measure linear density we use the nearest neighbour method²⁷, in which densities are estimated by taking the inverse of the width of the box required to contain k neighbouring points. The edges of sequential boxes meet between data points. Smoothing is controlled by k . We find that our results are constant for $k = 1$ –20, although the fitting error increases with k (Supplementary Fig. 9). We use $k = 1$, which produces the smallest error and minimum bias²⁷. The advantages of the nearest neighbour method are that the number and location of measurements are determined by the location of the data points, smoothing is uniform, and there are no empty bins.

The definition of linear density averages over all azimuths and is therefore insensitive to radiation pattern.

Catalogue completeness. We check catalogue completeness by fitting the Gutenberg-Richter magnitude frequency relationship²⁸, with $b = 1$. The aftershocks of $M = 2$ –3 mainshocks are complete to $M = 2$. For $M = 3$ –4 mainshocks, 10–15% of the smallest aftershocks ($M = 2.0$ –2.2) measured over the first 30 minutes may be missing, but there is no systematic loss with distance and thus no expected effect. Over the first 5 minutes, closer to 30–40% of the smallest aftershocks of the $M = 3$ –4 mainshocks may be missing, but again the loss is not significantly systematic with distance. Note that this incompleteness results in aftershock productivity appearing to vary as a factor of six with mainshock magnitude in Fig. 2, rather than the factor of ten noted in the literature^{20,24,25}. The $M \geq 3$ aftershocks used for the $M = 5$ –6 mainshocks are complete.

Goodness of fit. If the proposed single inverse power law fit is adequate, the residuals should have a distribution similar to that of data drawn from a pure power law distribution via Monte Carlo simulation. We test for similarity between the residuals of simulated data and the largest robust data set in this study (Southern California 30 min/16 km aftershocks of $M = 3$ –4 mainshocks) with the Kolmogorov-Smirnov test. At 95% and 65% confidence the test indicates that the null hypothesis—that the observed and simulated residuals come from the same distribution—cannot be rejected.

We test the composite power law $\rho(r) = a_1 r^{-n_1} + a_2 r^{-n_2}$ by fitting a_1 and a_2 with a nonlinear least-squares algorithm for every value of n_1 and n_2 between 0.5 and 3. We find that the single power law is preferred (bayesian information criterion).

Testing rate and state friction. To fit the rate and state friction equation to the aftershock density data, we try the parameters: $c_1 = 10^{-5.72}$ – $10^{-1.172}$; stress drop = 0.1–10 MPa; normal stress = 10–1,000 MPa; $A = 0.005$ –0.012 (ref. 3); $D = 0.1$ –2. The resulting range of c_2 is 0.008–200 $\times (M_0/4\pi)$. We achieve a minimum least-squared error at $D = 0.1$, $c_1 = 0.0034$ and $c_2 = 0.33(M_0/4\pi)$.

Potential hypocentre distribution effective dimension, D . Because of the limited seismogenic depth of Southern California, at length scales over ~ 10 km the system

would effectively be two-dimensional if potential hypocentres were randomly scattered throughout the crust. In reality, earthquakes concentrate on planar faults, whose width is also limited by the seismogenic depth. At distances longer than ~ 10 – 20 km, effective D for earthquakes randomly scattered on a fault tends towards 1. Multiple faults increase D , but the concentration of earthquakes in streaks and clusters on faults decreases it¹¹. Thus we expect $D \approx 1$ at large distances, and the lack of any break in the slope of the aftershock density decay, from 200 m to 100 km, for a wide range of mainshock magnitudes, suggests that $D \approx 1$ throughout.

Received 7 October 2005; accepted 10 April 2006.

- Gardner, J. K. & Knopoff, L. Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? *Bull. Seismol. Soc. Am.* **64**, 1363–1367 (1974).
- Stein, R. S. The role of stress transfer in earthquake occurrence. *Nature* **402**, 605–609 (1999).
- Dieterich, J. A. Constitutive law for the rate of earthquake production and its application to earthquake clustering. *J. Geophys. Res.* **99**, 2601–2618 (1994).
- Gomberg, J., Bodin, P. & Reasenber, P. A. Observing earthquakes triggered in the near field by dynamic deformations. *Bull. Seismol. Soc. Am.* **93**, 118–138 (2003).
- Johnson, P. A. & Jia, X. Nonlinear dynamics, granular media and dynamic earthquake triggering. *Nature* **437**, 871–874 (2005).
- Ichinose, G. A., Anderson, J. G. & Smith, K. D. Static stress change caused by the 1978 Diamond Valley, California and 1994 Double Spring Flat Nevada earthquakes. *Eos* **78**, abstr. S22B-04 (1997).
- Ogata, Y. Statistical models for earthquake occurrence and residual analysis for point processes. *J. Am. Stat. Assoc.* **83**, 9–27 (1988).
- Huc, M. & Main, I. G. Anomalous stress diffusion in earthquake triggering: Correlation length, time-dependence, and directionality. *J. Geophys. Res.* **108**, 2324, doi:10.1029/2001JB001645 (2003).
- Shearer, P., Hauksson, E. & Lin, G. Southern California hypocenter relocation with waveform cross-correlation, part 2: Results using source-specific station terms and cluster analysis. *Bull. Seismol. Soc. Am.* **95**, 904–915 (2005).
- Dodge, D. A., Beroza, G. C. & Ellsworth, W. L. Foreshock sequence of the 1992 Landers, California earthquake and its implications for earthquake nucleation. *J. Geophys. Res.* **100**, 9865–9880 (1995).
- Rubin, A. M., Gillard, D. & Got, J. Streaks of microearthquakes along creeping faults. *Nature* **400**, 635–641 (1999).
- Kagan, Y. Y. Short-term properties of earthquake catalogs and models of earthquake source. *Bull. Seismol. Soc. Am.* **94**, 1207–1228 (2004).
- Bak, P., Christensen, K., Danon, L. & Scanlon, T. Unified scaling law for earthquakes. *Phys. Rev. Lett.* **88**, 109901 (2002).
- Wells, D. L. & Coppersmith, K. J. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bull. Seismol. Soc. Am.* **84**, 974–1002 (1995).
- Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978).
- Main, I. G., Leonard, T., Papasouliotis, O., Hatton, C. G. & Meredith, P. G. One slope or two? detecting statistically significant breaks of slope in geophysical data, with application to fracture scaling relationships. *Geophys. Res. Lett.* **26**, 2801–2804 (1999).
- Vidale, J. E., Agnew, D. C., Johnston, M. J. S. & Oppenheimer, D. H. Absence of earthquake correlation with earth tides: An indication of high preseismic fault stress rate. *J. Geophys. Res.* **103**, 24567–24572 (1998).
- Omori, F. On the aftershocks of earthquakes. *J. Coll. Sci. Imp. Univ. Tokyo* **7**, 111–200 (1894).
- Kagan, Y. Y. & Knopoff, L. Spatial distribution of earthquakes: the two-point correlation function. *Geophys. J. R. Astron. Soc.* **62**, 303–320 (1980).
- Helmstetter, A., Kagan, Y. Y. & Jackson, D. D. Importance of small earthquakes for stress transfers and earthquake triggering. *J. Geophys. Res.* **110**, B05S08, doi:10.1029/2004JB003286 (2005).
- Campbell, K. W. *Strong Motion Attenuation Relations* (Academic, London, 2003).
- Richter, C. F. An instrumental earthquake-magnitude scale. *Bull. Seismol. Soc. Am.* **25**, 1–32 (1935).
- Kanamori, H. et al. Determination of earthquake energy release and M_L using terrascopes. *Bull. Seismol. Soc. Am.* **83**, 330–346 (1993).
- Michael, A. J. & Jones, L. M. Seismicity alert probabilities at Parkfield, California, revisited. *Bull. Seismol. Soc. Am.* **88**, 117–130 (1998).
- Felzer, K. R., Abercrombie, R. E. & Ekström, G. A common origin for aftershocks, foreshocks, and multiplets. *Bull. Seismol. Soc. Am.* **94**, 88–98 (2004).
- Davidsen, J. & Paczuski, M. Analysis of the spatial distribution between successive earthquakes. *Phys. Rev. Lett.* **94**, 048501 (2005).
- Silverman, B. W. *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, New York, 1986).
- Gutenberg, B. & Richter, C. F. Frequency of earthquakes in California. *Bull. Seismol. Soc. Am.* **4**, 185–188 (1944).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature

Acknowledgements We thank R. Abercrombie, A. Felzer, N. Field, M. Gerstenberger, J. Gomberg, S. Gross, J. Hardebeck, M. Harrington, A. Helmstetter, S. Hough, L. Jones, Y. Kagan, H. Kanamori, I. Main, S. Prejean, P. Shearer, R. Stein, J. Vidale, K. Richards-Dinger and A. Yong for comments. This work was supported in part by the National Science Foundation.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.R.F. (kfelzer@gps.caltech.edu).

Bone histology indicates insular dwarfism in a new Late Jurassic sauropod dinosaur

P. Martin Sander¹, Octávio Mateus², Thomas Laven³ & Nils Knötschke³

Sauropod dinosaurs were the largest animals ever to inhabit the land, with truly gigantic forms in at least three lineages^{1–3}. Small species with an adult body mass less than five tonnes are very rare^{4,5}, and small sauropod bones generally represent juveniles. Here we describe a new diminutive species of basal macronarian sauropod, *Europasaurus holgeri* gen. et sp. nov., and on the basis of bone histology we show it to have been a dwarf species. The fossils, including excellent skull material, come from Kimmeridgian marine beds of northern Germany^{6,7}, and record more than 11 individuals of sauropods 1.7 to 6.2 m in total body length. Morphological overlap between partial skeletons and isolated bones links all material to the same new taxon. Cortical histology of femora and tibiae indicates that size differences within the specimens are due to different ontogenetic stages, from juveniles to fully grown individuals. The little dinosaurs must have lived on one of the large islands around the Lower Saxony basin⁸. Comparison with the long-bone histology of large-bodied sauropods suggests that the island dwarf species evolved through a decrease in growth rate from its larger ancestor.

Sauropoda Marsh, 1878

Neosauropoda Bonaparte, 1886

Macronaria Wilson and Sereno, 1998

Europasaurus holgeri gen. et sp. nov.

Etymology. The generic name means ‘reptile from Europe’, after Europe and the Greek *sauros* for lizard; *holgeri* after Holger Lütke, who discovered the first bones.

Holotype. DFMMh/FV 291: disarticulated left premaxilla; right maxilla; right quadratojugal; occipital region; left laterosphenoid–orbitosphenoid complex; right surangular; right angular; left dentary; teeth; cervical and sacral vertebrae; and cervical and dorsal ribs of one individual (see Fig. 1 and Supplementary Information). DFMMh/FV: Dinosaurier-Freilichtmuseum Münchenhagen/Verein zur Förderung der Niedersächsischen Paläontologie (e.V.), Germany. **Referred material.** Cranial and postcranial elements of at least ten individuals, preserved as isolated bones to partially articulated skeletons, including young juveniles (estimated body length 1.7 m) to adults (body length 6.2 m).

Horizon and locality. Late Jurassic, middle Kimmeridgian marine carbonate rock, bed 93 of section at Langenberg quarry⁷, Lower Saxony basin, Oker near Goslar, Niedersachsen, northern Germany (see Supplementary Information).

Diagnosis. *Europasaurus holgeri* gen. et sp. nov. shows the following unambiguous autapomorphies (see also Supplementary Information): nasal process of premaxillary projecting anterodorsally; medial notch on posterior dorsal margin of cervical vertebral centra; scapular acromion with a prominent posterior projection; and transverse width of astragalus twice its dorsoventral height and anteroposterior length. *Europasaurus* differs from *Camarasaurus* in

the wing-shaped posterior process of the postorbital being slightly longer and wider than the anterior process, whereas it is much shorter in *Camarasaurus*. *Europasaurus* also differs from *Camarasaurus* in its short nasal–frontal contact; rectangular parietal in posterior view; and undivided presacral neural spines. *Europasaurus* differs from *Brachiosaurus* in the shorter muzzle, quadratojugal contacting squamosal; participation of jugal in ventral margin of skull; short nasal–frontal contact; and humerus flat anteromedially with proximal and distal epiphyses not aligned. *Europasaurus* differs from *Lusotitan atalaiensis*²⁵ in the shape of the ilium and the astragalus, and from the potentially valid ‘*Cetiosaurus*’ *humero cristatus*⁴ in its shorter and less prominent deltopectoral crest. *Europasaurus* differs from almost all known neosauropods in its diminutive adult body size.

Phylogenetic analysis (see Supplementary Information) indicates that *Europasaurus holgeri* gen. et sp. nov. is a macronarian that is more derived than *Camarasaurus*, and is the sistergroup of Brachiosauridae and all (more-derived) Titanosauromorpha. It also indicates that the diminutive body size of *Europasaurus* is derived.

Six individuals that represent the full body-length range known for *Europasaurus* were sampled histologically from one or two long bones each. The bones selected were femora and tibiae, and the bone tissues examined were those of the cortex (see Supplementary Information). The bone cortex of the smallest individual (body length 1.75 m; DFMMh/FV 009) is primary bone of the fibrolamellar complex with a reticular vascular network that grades into a laminar network (Fig. 2a). The bone matrix consists of fibrous tissue with plump osteocyte lacunae. Only a thin veneer of lamellar bone lines the vascular canals (incipient primary osteons). There are no growth marks in the cortical bone. The inner cortex lacks any indication of secondary remodelling, except for large erosional cavities. These features indicate a rapidly growing juvenile^{9–11}.

The next-largest individual (body length 2 m; DFMMh/FV 291.9) differs from the previous one in that the fibrolamellar bone is of the laminar type, the vascular network being organized into a predominantly circumferential pattern. The vascular canals have a lining of lamellar bone, forming primary osteons. In the next-largest individual (body length 3.5 m; DFMMh/FV 001), the primary osteons in the fibrolamellar complex are mature with a narrow central vascular canal. Notably, there are two cyclical growth marks (annuli) in the fibrolamellar bone of the outer cortex.

A slightly larger individual (body length 3.7 m; DFMMh/FV 495) was sampled from its tibia and femur. Both bones have the same histology of laminar fibrolamellar bone interrupted by growth marks (Fig. 2b), the spacing of which diminishes towards the outer bone surface, indicating a decrease in growth rate. However, the vascular canals in the outermost cortex are large and open to the bone surface, indicating active growth at the time of death. The next-largest individual (body length 4.6 m; DFMMh/FV 153) has the same primary bone histology as the previous ones, preserving five growth

¹Institute of Paleontology, University of Bonn, Nussallee 8, D-53115 Bonn, Germany. ²Centro de Estudos Geológicos da Universidade Nova de Lisboa and Museu da Lourinhã, Rua João Luis de Moura, 2530-157 Lourinhã, Portugal. ³Dinosaurier-Freilichtmuseum Münchenhagen, Alte Zollstrasse 5, D-31547 Rehburg-Loccum, Germany.

marks with decreasing spacing. It differs in having isolated secondary osteons in the inner cortex.

The largest individual (body length 6.2 m; DFMMh/FV 415) was sampled from a distal femur fragment. Because of the distal position of the sample, the cortex is relatively thinner than in the other specimens (Fig. 2c). In the inner part, the laminar fibrolamellar bone is heavily remodelled by secondary osteons with little primary bone remaining. The outer cortex has three closely spaced lines of arrested growth (LAGs). Vascular canals are longitudinal, and the tissue is predominantly lamellar. Vascularity decreases outwards, with the last cycle being nearly avascular and consisting of lamellar bone only. This histology is typical of an external fundamental system (ESF), and indicates that the individual was fully grown^{9,11–13}. Because only one cycle of purely lamellar bone was laid down, the animal must have died soon after reaching full size.

Four size-related trends in the cortical histology of the *Europasaurus* long bones indicate that the fossils represent a growth series from juveniles to subadults to fully grown adults^{9,11,12}: (1) only medium-sized and larger individuals have growth marks (annuli and LAGs) that diminish in spacing towards the outer surface, recording a decreasing growth rate; (2) only the largest individual was fully grown; (3) vascularization decreases from the smallest to the largest individuals, which is coupled with increased organization of the vascular network; and (4) haversian remodelling increases with

increasing body size, from none in the four smallest specimens to many secondary osteons in the largest.

Palaeogeography⁸ suggests insular dwarfing as the explanation for the diminutive body size of *Europasaurus*. The largest palaeo-islands surrounding the locality in the Lower Saxony basin had areas of <200,000 km² (calculated from ref. 8). Such islands would not have been able to support large-bodied sauropods. The ancestor of *Europasaurus* would have dwarfed rapidly on immigrating to the island, or as a response to shrinking land masses caused by rising sea levels. Previous hypotheses about island dwarfs among dinosaurs have focused on the Cretaceous of southeastern Europe^{14–17}, in particular on the sauropod *Magyarosaurus dacus* and the hadrosaur *Telmatosaurus transylvanicus* from the latest Cretaceous of the Hateg basin in Romania^{14,15,18}. Fusion of cranial bones suggests that *Telmatosaurus* was fully grown¹⁸, but the ontogenetic status of *Magyarosaurus* remains uncertain^{16,19} in the absence of histological or bone fusion evidence.

The development of growth marks in the long bone cortex of *Europasaurus* suggests that it grew more slowly than larger less-derived neosauropod dinosaurs, including *Camarasaurus*^{12,20–22}. Thus, an evolutionary decrease in growth rate was important in the dwarfing of *Europasaurus*. This is a reversal of the accelerated growth in the evolution of giant body size in sauropod and theropod dinosaurs^{20,23,24}.

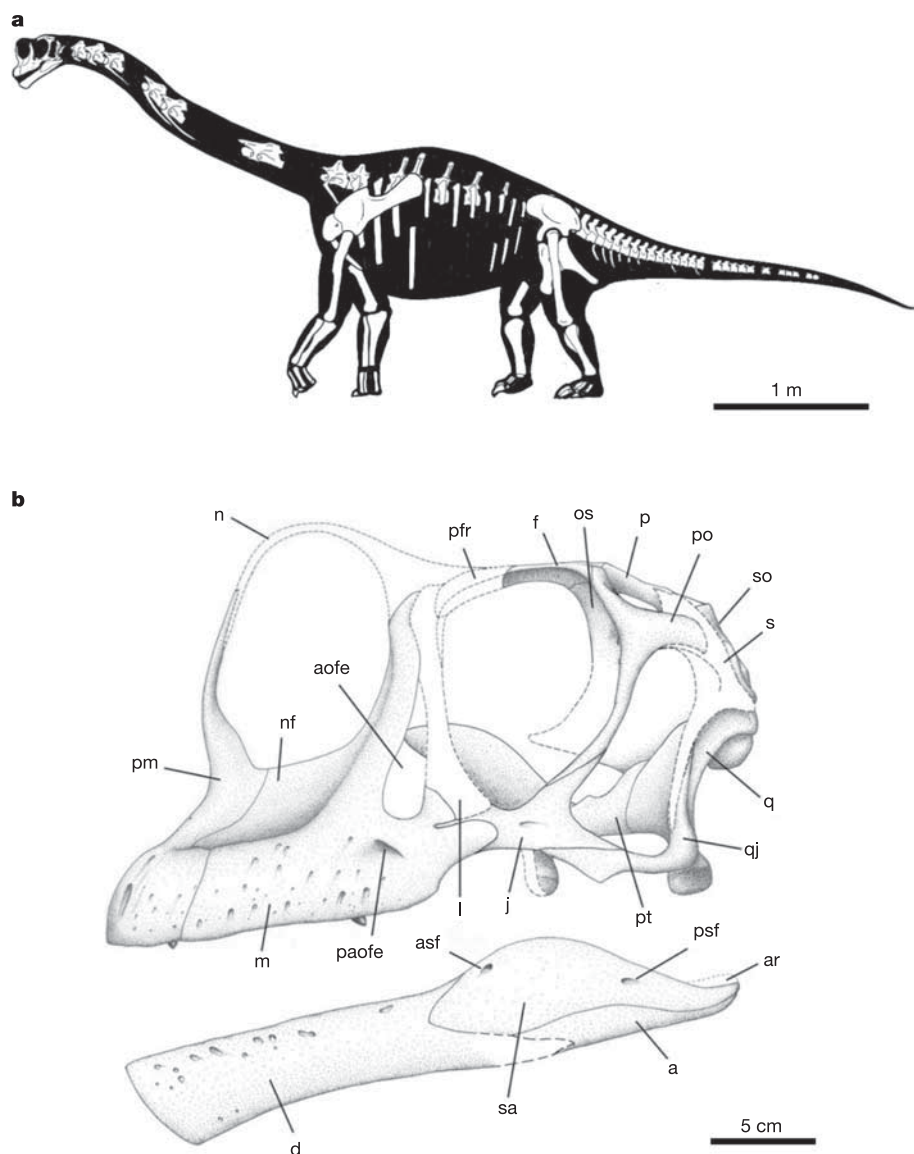


Figure 1 | *Europasaurus holgeri* gen. et sp. nov. **a**, Reconstruction of the skeleton (known parts shown in white) and outline (black) scaled to the largest individual (body length 6.2 m). **b**, Skull reconstruction based on several individuals but scaled to the holotype. Abbreviations: a, angular; aofe, antorbital fenestra; ar, articular; asf, anterior surangular foramen; d, dentary; f, frontal; j, jugal; l, lacrimal; m, maxilla; n, nasal; nf, narial fossa; os, orbitosphenoid; p, parietal; paofe, preantorbital fenestra; pfr, prefrontal; pm, premaxillary; po, postorbital; psf, posterior surangular foramen; pt, pterygoid; q, quadrate; qj, quadratojugal; s, squamosal; sa, surangular; so, supraoccipital.

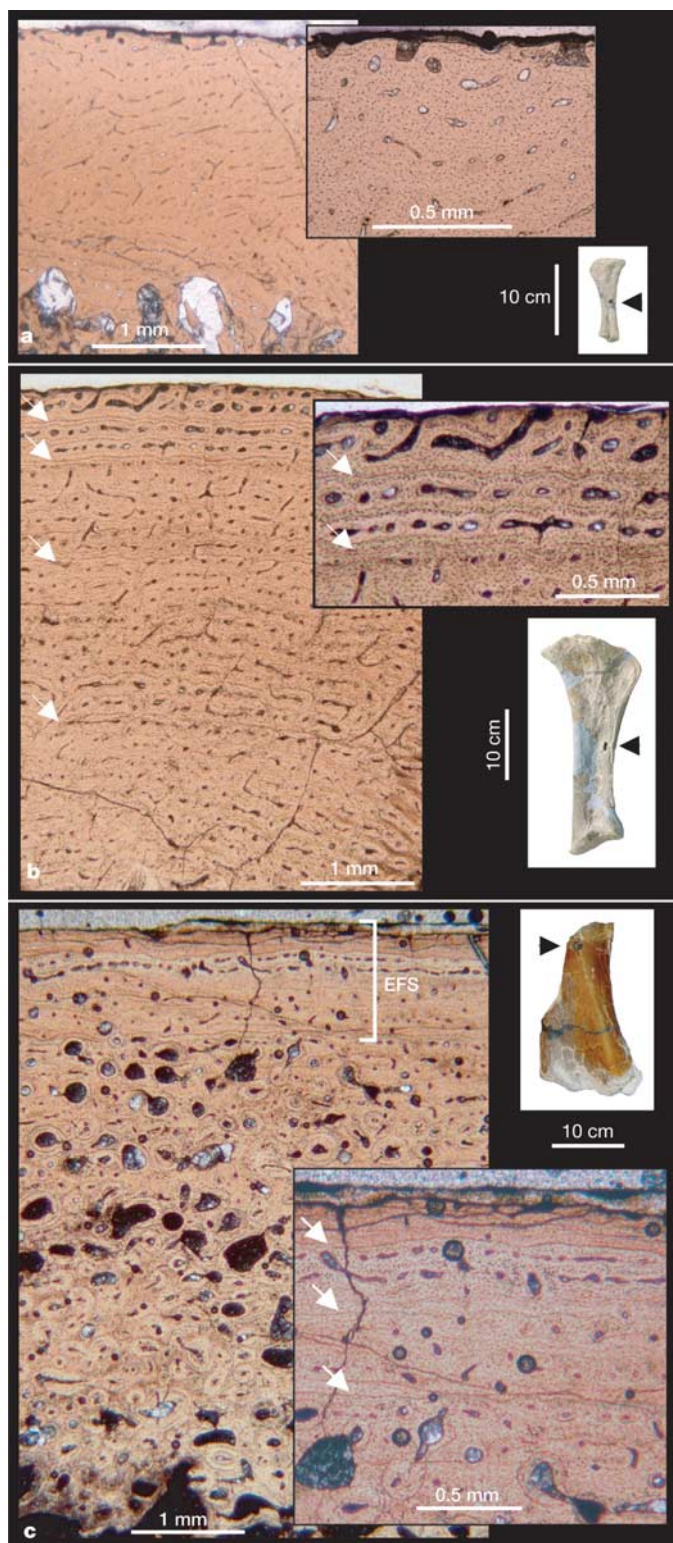


Figure 2 | Histological growth series and sampled bones of *Europasaurus holgeri* gen. et sp. nov. **a**, Tibia from DFMMh/FV 009, the smallest individual (body length 1.75 m). The reticular fibrolamellar tissue, which grades into laminar fibrolamellar tissue (inset), and the absence of growth marks indicate its juvenile status. **b**, Tibia from DFMMh/FV 459.5, a mid-sized individual (body length 3.7 m). The cortex consists of laminar fibrolamellar bone interrupted by growth marks (arrows). Wide vascular canals opening to the outer bone surface (inset) indicate that this animal was still actively growing. **c**, Distal femur from DFMMh/FV 415, the largest individual (body length 6.2 m). The external fundamental system (EFS; inset) indicates that it was fully grown. Bone surface is at the top of all photomicrographs. Black arrows indicate sample locations; white arrows indicate growth marks.

Received 20 December 2005; accepted 7 February 2006.

1. Wilson, J. A. Sauropod dinosaur phylogeny: critique and cladistic analysis. *Zool. J. Linn. Soc.* **136**, 217–276 (2002).
2. Wedel, M. J., Cifelli, R. L. & Sanders, R. K. *Sauroposeidon proteles*, a new sauropod from the Early Cretaceous of Oklahoma. *J. Vertebr. Paleontol.* **20**, 109–114 (2000).
3. Smith, J. B. et al. A giant sauropod dinosaur from an Upper Cretaceous mangrove deposit in Egypt. *Science* **292**, 1704–1706 (2001).
4. Upchurch, P., Barrett, P. M. & Dodson, P. in *The Dinosauria* 2nd edn (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 259–322 (Univ. of California Press, Berkeley, 2004).
5. Rauhut, O. W. M. et al. Discovery of a short-necked sauropod dinosaur from the Late Jurassic period of Patagonia. *Nature* **435**, 670–672 (2005).
6. Lotze, F. Zum Jura des Langenberges zwischen Oker und Bad Harzburg (nördl. Harzrand). *Neues Jb. Geol. Paläont. Mh.* **1968**, 730–732 (1968).
7. Pape, H. Die Malmsschichtenfolge vom Langenberg bei Oker (nördl. Harzvorland). *Mitt. Geol. Inst. Tech. Univ. Hann.* **9**, 41–134 (1970).
8. Ziegler, P. A. *Geological Atlas of Western and Central Europe* 2nd edn (Shell International Petroleum Company, Amsterdam, 1990).
9. Sander, P. M. Long bone histology of the Tendaguru sauropods: implications for growth and biology. *Paleobiology* **26**, 466–488 (2000).
10. Horner, J. R., Padian, K. & de Ricqlès, A. Comparative osteohistology of some embryonic and neonatal archosaurs: implications for variable life histories among dinosaurs. *Paleobiology* **27**, 39–58 (2001).
11. Erickson, G. M. Assessing dinosaur growth patterns: a microscopic revolution. *Trends Ecol. Evol.* **20**, 677–684 (2005).
12. Chinsamy-Turan, A. *The Microstructure of Dinosaur Bone* (Johns Hopkins Univ. Press, Baltimore, 2005).
13. Turvey, S. P., Green, O. R. & Holdaway, R. N. Cortical growth marks reveal extended juvenile development in New Zealand moa. *Nature* **435**, 940–943 (2005).
14. Nopcsa, F. Über das Vorkommen der Dinosaurier in Siebenbürgen. *Verh. Zool. Bot. Ges. Wien* **54**, 12–14 (1914).
15. Weishampel, D., Grigorescu, D. & Norman, D. B. The dinosaurs of Transylvania. *Natl. Geogr. Res.* **7**, 196–215 (1991).
16. Jianu, C. M. & Weishampel, D. B. The smallest of the largest: a new look at possible dwarfing in sauropod dinosaurs. *Geol. Mijnbouw* **78**, 335–343 (1999).
17. Dalla Vecchia, F. M. et al. Dinosaur track sites in the upper Cenomanian (Late Cretaceous) of the Istrian Peninsula (Croatia). *Bol. Soc. Paleont. Ital.* **40**, 25–53 (2001).
18. Weishampel, D. B., Norman, D. B. & Grigorescu, D. *Telmatosaurus transylvanicus* from the Late Cretaceous of Romania: the most basal hadrosaurid dinosaur. *Palaeontology* **36**, 361–385 (1993).
19. Le Loeuff, J. Romanian Late Cretaceous dinosaurs: big dwarfs or small giants? *Hist. Biol.* **17**, 15–17 (2005).
20. Padian, K., de Ricqlès, A. & Horner, J. R. Dinosaurian growth rates and bird origins. *Nature* **412**, 405–408 (2001).
21. de Ricqlès, A. et al. in *Bone Vol. 3 Bone Matrix and Bone Specific Products* (ed. Hall, B. K.) 1–78 (CRC, Boca Raton, 1991).
22. Sander, P. M. & Tückmantel, C. Bone lamina thickness, bone apposition rates, and age estimates in sauropod humeri and femora. *Paläontol. Z.* **76**, 161–172 (2003).
23. Erickson, G. M. et al. Gigantism and comparative life-history parameters of tyrannosaurid dinosaurs. *Nature* **430**, 772–774 (2004).
24. Sander, P. M. et al. Adaptive radiation in sauropod dinosaurs: bone histology indicates rapid evolution of giant body size through acceleration. *Org. Divers. Evol.* **4**, 165–173 (2004).
25. Antunes, M. T. & Mateus, O. Dinosaurs of Portugal. *C. R. Palevol.* **2**, 77–95 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank O. Dülfer, W. Fink, D. Flemming, W. Fricke, A. Hänel, O. Heumann, D. Kranz, M. Lorsch, M. Mastroianni, R. Nimser, G. Oleschinski, U. Resch, D. Rössler, H.-J. Siber, S. Thiele, D. Unwin, F. Wesling, J. Wilson and R. Windolf. We especially appreciate the support of J. McIntosh, F. von Pupka, B. Wolter and H. Lüdte. This research was supported by the Jurassic Foundation, the Deutsche Forschungsgemeinschaft (DFG) and the Portuguese “Fundação para a Ciência e Tecnologia”. This paper is contribution number 9 of the DFG Research Unit 533 “Biology of the Sauropod Dinosaurs: The Evolution of Gigantism”.

Author Contributions P.M.S. was responsible for the bone histology work presented as part of this study. The morphology, systematics and taphonomy of the new sauropod was studied by the remaining authors, who are to be considered the sole authors of the name *Europasaurus holgeri* gen. et sp. nov.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.M.S. (martin.sander@uni-bonn.de) or O.M. (omateus@museulourinha.org).

LETTERS

Intensity of sexual selection along the anisogamy–isogamy continuum

Adam Bjork¹ & Scott Pitnick¹

Research into the evolution of giant sperm has uncovered a paradox within the foundations of sexual selection theory. Postcopulatory sexual selection on males (that is, sperm competition and cryptic female choice) can lead to decreased sperm numbers by favouring the production of larger sperm¹. However, a decline in sperm numbers is predicted to weaken selection on males and increase selection on females^{2,3}. As isogamy is approached (that is, as investment per gamete by males approaches that by females), sperm become less abundant, ova become relatively less rare, and competition between males for fertilization success is predicted to weaken. Sexual selection for longer sperm, therefore, is expected to be self limiting. Here we examine this paradox in *Drosophila* along the anisogamy–isogamy continuum using intraspecific experimental evolution techniques and interspecific comparative techniques. Our results confirm the big-sperm paradox by showing that the sex difference in sexual selection gradients⁴ decreases as sperm size increases. However, a resolution to the paradox is provided when this finding is interpreted in concert with the ‘opportunity for selection’ and the ‘opportunity for sexual selection’^{5,6}. Furthermore, we show that most of the variation in measures of selection intensity is explained by sperm length and relative investment in sperm production.

Bateman’s² quantitative description of sex differences in *Drosophila melanogaster* gave rise to the modern era of sexual selection theory^{3,7–9} by showing that the slope of the line relating reproductive success to mating success (the sexual selection gradient) is nearly flat for females, whereas the slope of this line is much steeper for males. The magnitude of the sex difference in the strength of selection depends on the relationship between male and female sexual selection gradients^{4,10,11}. Anisogamy generates the conditions for sexual selection, as numerically abundant male gametes compete to fertilize rare female gametes¹².

In contrast, in a truly ‘isogamous’ population, where males and females produce identical numbers of equal-sized gametes, it would be possible for every gamete to participate in a successful fertilization. Male and female sexual selection gradients would converge and have equivalent slopes, and the intensity of selection on each sex would be identical, assuming no parental care¹³. This theory has not been tested, however, because exceptionally high ratios of sperm number to egg number have been considered ubiquitous across taxa.

Selection generated by sperm competition is attributed with the evolutionary maintenance of anisogamy, or tiny sperm^{14–16}. However, recent comparative analyses of some taxa^{17,18}, and experimental evolution studies in *D. melanogaster*¹ and *Caenorhabditis elegans*¹⁹, indicate that postcopulatory sexual selection can also favour increased investment per sperm. The rise in costs associated with the production of longer sperm—including delayed reproductive maturity²⁰, decreased male fecundity^{20,21} and increased energetic investment in testes (as measured by the gonadosomatic index, or GSI = (gonad mass/total body mass) × 100)—suggests that the

strength of selection maintaining sperm length does not decline as isogamy is approached. Evidence suggesting intense sexual selection in species with longer sperm conflicts with the theoretical prediction that sexual selection should be weaker in species with longer sperm owing to the trade-off with sperm number.

Substantial variation in sperm size has been described for most taxa²²; sperm length across *Drosophila* species is more variable than in the remainder of the animal kingdom²³. Therefore, this genus serves as a useful system to examine the big-sperm paradox. In species with no parental care, as in most *Drosophila*²⁴, parental investment consists of the energy invested in sperm or eggs, and potential reproductive rates⁸ can be determined by measuring gamete production rates. We quantified sex-specific gamete production rates in both a short- and long-sperm species: *D. melanogaster* (sperm length = 1.87 mm; GSI = 5.05; ref. 21) and *D. bifurca* (sperm length = 58.29 mm; GSI = 10.60; ref. 25). A sperm:egg production rate ratio of 29.3:1 for *D. melanogaster* versus 5.8:1 for *D. bifurca* was determined (Table 1). *D. bifurca* is nearly isogamous in terms of gamete size and gamete production rate (Fig. 1). Natural populations of *D. bifurca* probably further approach isogamy because males require 17 days posteclosion to produce fertile gametes, whereas females require only 7 days (ref. 25)—a factor not included in our rate calculations. GSI, which is widely reported and easily quantifiable, was used in the among-species analyses. Sperm length in *Drosophila* explains nearly all of the interspecific variation in GSI²¹. Although interspecific relationships are reported here using only GSI, qualitatively similar findings resulted from analyses using sperm length, and multiple regressions confirmed that testis mass—as opposed to body mass—is the component of GSI that explains the majority of the variation among species in all significant correlations (see below).

We repeated Bateman’s² competitive mating experiment along the anisogamy–isogamy continuum with experimental evolution lines of *D. melanogaster* selected for longer (mean ± s.e.m. = 2.03 ± 0.01 mm) or shorter (mean ± s.e.m. = 1.67 ± 0.01 mm; ref. 1) sperm lengths. The resulting sexual selection gradients confirm the big-sperm paradox: as sperm length increases, the magnitude of sex differences declines (Fig. 3a). This decline results from the combined effect of a decrease in male slope (analysis of covariance (ANCOVA), $F_{1,58} = 4.689$, $P = 0.0345$) and, although not significant, an increase in female slope (ANCOVA, $F_{1,244} = 0.340$, $P = 0.5605$) as sperm size increases (Fig. 2a, b).

We continued our investigation interspecifically by conducting the ‘Bateman experiment’ with a separate, non-experimentally evolved *D. melanogaster* population, *D. bifurca*, and two species (*D. virilis* and *D. lummei*) with intermediate sperm lengths (5.70 mm and 7.79 mm, respectively²⁰) and intermediate GSI (5.79 and 8.04, respectively; S.P., unpublished data). These experiments (Fig. 2c–f) showed that the sex differences in sexual selection gradients share the same negative relationship with investment in sperm production at the

¹Department of Biology, Syracuse University, Syracuse, New York 13244-1270, USA.

Table 1 | Sperm and egg production rates of *D. melanogaster* and *D. bifurca*

Species	Sperm length (mm)	GSI*	Gamete production rate†				Sperm:egg production rate
			Female		Male		
			No. eggs per day	No. sperm (t = 0 h)	No. sperm (t = 6 h)	No. sperm per day	
<i>D. melanogaster</i>	1.87	5.05	59.46 ± 1.80 (n = 45)	806.00 ± 99.72 (n = 8)	1,242.25 ± 144.59 (n = 8)	1,745	29.3:1
<i>D. bifurca</i>	58.29	10.60	38.13 ± 2.43 (n = 39)	78.36 ± 8.39 (n = 11)	133.64 ± 8.63 (n = 11)	221	5.8:1

*GSI, gonadosomatic index.

†Values are mean ± s.e.m.

macroevolutionary level as they do intraspecifically (Fig. 3b). The slope for males is significantly steeper than the slope for females in *D. melanogaster* (Fig. 2c) and *D. virilis* (Fig. 2d). In contrast, sex-specific slopes of *D. lummei* (Fig. 2e) and *D. bifurca* (Fig. 2f) are not statistically different. When comparing across all species, we discovered that male GSI explains most of the variation (that is, 93.5%) in sex difference in selection gradients (Fig. 3b).

The 'opportunity for sexual selection' (I_s ; ref. 5) is a standardized index—based on variances in reproductive success—of sexual selection intensity on males and the sex difference in the strength of selection⁹. I_s is determined by subtracting the 'opportunity for selection' ($I = \text{variance in reproductive success} / (\text{mean reproductive success})^2$; ref. 5) for females (I_{females}) from that for males (I_{males}). I_s and sexual selection gradients were expected to complement each other because the greater the male selection gradient slope, relative to the female slope, the greater the expected intensity of intramale competition for mates. However, I_s did not decrease across species—and, in fact, it increased significantly within species—as sperm size increased (Fig. 3d and c, respectively). Moreover, this positive relationship with sperm length was detected within species for both I_{males} and I_{females} (Fig. 3e, g) and among species for I_{females} (GSI explained 99.4% of the variation; Fig. 3h).

Why do the patterns revealed from analyses of sexual selection gradients and I_s differ? Although both approaches measure selection, they measure different aspects of the process. I_s estimates the overall intensity of sexual selection. Sex differences in selection gradients, however, measure the degree to which that selection may operate differentially on the sexes.

The resolution of the big-sperm paradox is achieved when sexual selection gradients and I_s are interpreted in tandem. In the most anisogamous species examined, *D. melanogaster*, I_s is relatively small. However, the disparity in slope between the male and female

selection gradients demonstrates that selection on male, but not female, mating competition is likely to be a strong force in this species because increased male mating success leads to markedly improved reproductive success²⁶. High I_s in *D. bifurca* and *D. lummei* exists despite there being no significant difference between the male and female sexual selection gradients within these species. This bolsters the claims of recent empirical work that sperm gigantism in *Drosophila* is a product of intense sexual selection^{1,21}. Historically, models of sex differences have considered the evolution of sperm size strictly from the perspective of initial parental investment^{3,15}. We contend that exaggerated sperm tails should not be considered as a form of parental investment in offspring or as a material gift to females. For example, in *D. bifurca* only a tiny portion of the sperm enters the egg; the vast majority is used neither by the egg nor the female²⁷. These long sperm flagella are best thought of as ornaments or armaments—the result of directional postcopulatory sexual selection for traits that enhance competitive fertilization success¹.

The trade-off between sperm size and sperm number probably contributes to the significant intra- and inter-specific increases reported for I_{females} as sperm size increases (Fig. 3g, h). Higher variance in female reproductive success is expected if fewer sperm are available and the sperm storage organs of all females are not filled to capacity. This, in turn, will lead to selection for increased female re-mating to ensure fertilization in systems with longer sperm. This aspect of our study will be examined in detail in subsequent work, with attention given to the fitness costs associated with multiple mating in females, optimal female re-mating rate, and the inter-relationship of female re-mating rate and sperm size with the strength of sexual selection at the precopulatory versus postcopulatory stage.

The solving of the big-sperm paradox provides a fresh perspective on sexual selection theory by focusing attention on the dual function

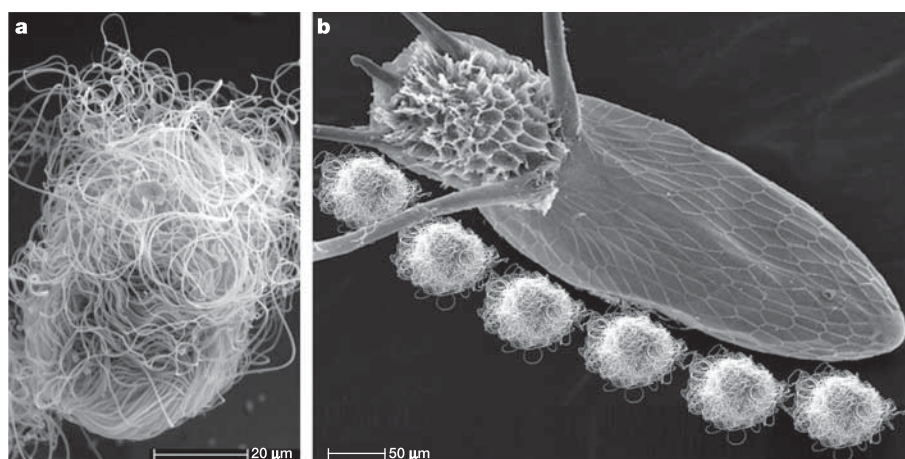


Figure 1 | *Drosophila bifurca* sperm and egg. Sperm are produced at a rate that is approximately six times faster than eggs in *D. bifurca* (see Table 1). **a**, Scanning electron micrograph (SEM) showing a single, 6-cm *D. bifurca* spermatozoon dissected from the seminal vesicle, where sperm are

individually rolled into compact balls. **b**, SEM of a single *D. bifurca* sperm (copied six times) next to an SEM of a *D. bifurca* ovum at the same magnification. Micrographs by R. Dallai.

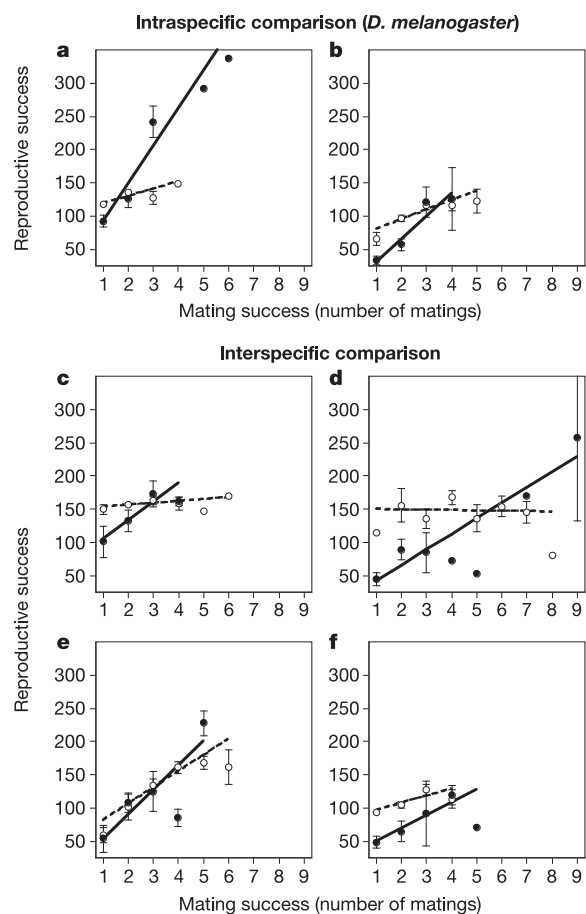


Figure 2 | Intraspecific and interspecific sexual selection gradients in *Drosophila*. Open circles (dashed lines) represent females; closed circles (solid lines) represent males. Each symbol represents a mean, though regressions (see Supplementary Table 1) are based on the complete raw data sets; error bars represent one s.e.m. Significance tests refer to analyses of covariance (ANCOVA) showing that male and female slopes are statistically different in all species/populations except *D. bifurca* and *D. lummei*, which manufacture the largest sperm. **a**, *D. melanogaster* short-sperm population, $F_{1,151} = 36.511$, $P < 0.0001$. **b**, *D. melanogaster* long-sperm population, $F_{1,151} = 6.119$, $P = 0.0145$. **c**, *D. melanogaster*, $F_{1,172} = 9.237$, $P = 0.0027$. **d**, *D. virilis*, $F_{1,52} = 13.972$, $P = 0.0005$. **e**, *D. lummei*, $F_{1,99} = 1.745$, $P = 0.1896$. **f**, *D. bifurca*, $F_{1,135} = 0.922$, $P = 0.3388$.

of sperm as both primary and secondary sexual traits. Our results underscore the importance of considering sex-specific gamete investment strategies²² and postcopulatory processes when exploring the nature of sex differences. The joint analysis of I_s and sexual selection gradients provides a resolution to the paradox. Previously, it has been widely recognized that female re-mating rate, parental investment, and operational sex ratio are critical descriptors of mating systems and sex differences^{3,7,8,28}. This current study suggests that sperm length and relative investment in sperm production serve as additional indicators of the most widely accepted measures of sexual selection intensity. Thus, sperm size and spermatogenic investment may provide simple and accurate assays for comparative analyses of the strength of sexual selection in the vast number of species without postmating parental investment.

METHODS

Experimental animals. *D. virilis* and *D. lummei* were obtained from the Tucson Drosophila Species Stock Center. The *D. bifurca* strain was derived from individuals collected near San Luis Potosi (SLP), Mexico in June 2002. For the intraspecific experimental analyses, we used lines of *D. melanogaster* subjected to bidirectional artificial selection for sperm length (see ref. 1). For the interspecific

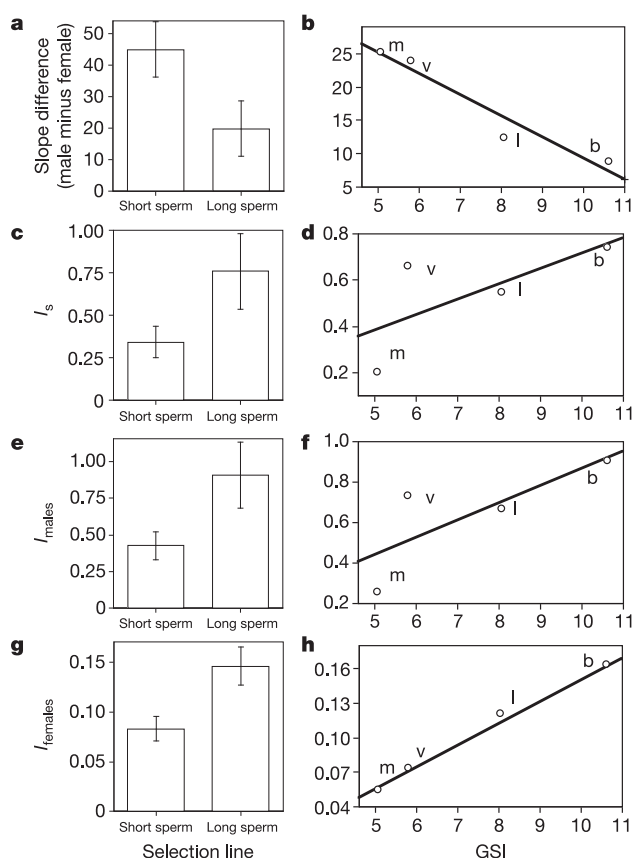


Figure 3 | Mating system measures in relation to investment in sperm production. Each interspecific regression point (open circles) represents a *Drosophila* species ($m = D. melanogaster$, $v = D. virilis$, $l = D. lummei$, $b = D. bifurca$). Error bars, which represent one s.e.m., and significance values for comparisons of the *D. melanogaster* selection lines were obtained using the bootstrap method with replacement (number of replications = 1,000). **a**, Intraspecifically, the sex difference in selection gradients is less in the long-sperm population than in the short-sperm population ($P = 0.018$). **b**, Sex difference in sexual selection gradients decreases significantly across species as investment in testes (GSI) increases ($y = -3.189x + 41.122$, $F_{1,2} = 28.807$, $R^2 = 0.935$, $P = 0.0330$, $n = 4$). **c**, In contrast, the opportunity for sexual selection, I_s , is greatest in the long-sperm *D. melanogaster* population ($P = 0.055$). **d**, I_s increases with GSI across species, though not significantly ($y = 0.067x + 0.084$, $F_{1,2} = 1.942$, $R^2 = 0.493$, $P = 0.2981$, $n = 4$). **e**, The opportunity for selection on males, I_{males} , is marginally greater for long-sperm males ($P = 0.068$). **f**, Interspecific changes in I_{males} do not correlate with GSI ($y = 0.086x + 0.008$, $F_{1,2} = 3.187$, $R^2 = 0.614$, $P = 0.2162$, $n = 4$). **g**, $I_{females}$ is greater in the long-sperm population ($P = 0.006$). **h**, $I_{females}$ rises significantly with GSI ($y = 0.019x - 0.04$, $F_{1,2} = 322.259$, $R^2 = 0.994$, $P = 0.0031$, $n = 4$).

comparative analyses, experiments with *D. melanogaster* used a large outbred population that had adapted to the laboratory for over 250 generations (LH_m; provided by A. Chippindale). Additional rearing conditions are described in Supplementary Information.

Gamete production rate. Methods used to quantify sperm and egg production rates are described in Supplementary Information.

Intensity of sexual selection. Copulations were observed in competitive mating vials containing four males and four females for 4 h each morning over four consecutive days. To identify individuals in copulating pairs, wings were uniquely clipped in a manner that does not affect mating success²⁹, defined as the copulation number. Flies were separated into individual vials for the remaining 20 h of each day. Female reproductive success was determined by counting all eggs laid in these vials. We used the sterile-male technique to determine male reproductive success. Within each mating vial, one of the four males (the 'focal male') was wing-clipped. The other three males were exposed to an X-ray dose (15 krad for *D. melanogaster*; 17.5 krad for other species) determined previously to sterilize sperm without disrupting fertilization (that

is, eggs were fertilized but did not hatch). Male mating success and reproductive success, determined by counting the number of hatched eggs laid by their mates, was quantified for focal males only. Individuals that failed to mate were included in the calculation of I_s . Because the most salient aspect of the sexual-selection-gradient approach addresses the sex-specific fecundity benefits of multiple mating, we chose not to include the zero-mating category in the analyses of sexual selection gradients. However, its inclusion produced qualitatively identical results. Supplementary Information provides a more detailed description of the analyses.

Scanning electron microscopy. The scanning electron microscopy (SEM) procedure is described in Supplementary Information.

Received 26 January; accepted 27 February 2006.

1. Miller, G. T. & Pitnick, S. Sperm–female coevolution in *Drosophila*. *Science* **298**, 1230–1233 (2002).
2. Bateman, A. J. Intrasexual selection in *Drosophila*. *Heredity* **2**, 349–368 (1948).
3. Trivers, R. L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B.) 52–97 (Aldine Publishing, Chicago, 1972).
4. Arnold, S. J. Bateman's principals and the measurement of sexual selection in plants and animals. *Am. Nat.* **144**, s126–s149 (1994).
5. Wade, M. J. Sexual selection and variance in reproductive success. *Am. Nat.* **114**, 742–747 (1979).
6. Wade, M. J. & Arnold, S. J. The intensity of sexual selection in relation to male sexual behaviour, female choice, and sperm precedence. *Anim. Behav.* **28**, 446–461 (1980).
7. Emlen, S. & Oring, L. W. Ecology, sexual selection, and the evolution of mating systems. *Science* **197**, 215–223 (1977).
8. Clutton-Brock, T. H. & Parker, G. A. Potential reproductive rates and the operation of sexual selection. *Q. Rev. Biol.* **67**, 437–456 (1992).
9. Shuster, S. M. & Wade, M. J. *Mating Systems and Strategies* (Princeton Univ. Press, Princeton, New Jersey, 2003).
10. Jones, A. G., Rosenqvist, G., Anders, B., Arnold, S. J. & Avise, J. C. The Bateman gradient and the cause of sexual selection in a sex-role-reversed pipefish. *Proc. R. Soc. Lond. B* **267**, 677–680 (2000).
11. Jones, A. G., Arguello, J. R. & Arnold, S. J. Validation of Bateman's principles: a genetic study of mating patterns and sexual selection in newts. *Proc. R. Soc. Lond. B* **269**, 2533–2539 (2002).
12. Kokko, H. & Jennions, M. It takes two to tango. *Trends Ecol. Evol.* **18**, 103–104 (2003).
13. Lorch, P. D. Understanding reversals in the relative strength of sexual selection on males and females: a role for sperm competition? *Am. Nat.* **159**, 645–657 (2002).
14. Parker, G. A. Sperm competition and its evolutionary consequences in insects. *Biol. Rev.* **45**, 525–567 (1970).
15. Parker, G. A. Why are there so many tiny sperm? Sperm competition and the maintenance of two sexes. *J. Theor. Biol.* **96**, 281–294 (1982).
16. Parker, G. A., Baker, R. R. & Smith, V. G. F. The origin of evolution of gamete dimorphism and the male–female phenomenon. *J. Theor. Biol.* **36**, 529–553 (1972).
17. Briskie, J. V. & Montgomerie, R. Sperm size and sperm competition in birds. *Proc. R. Soc. Lond. B* **247**, 89–95 (1992).
18. Gage, M. J. G. Associations between body size, mating pattern, testis size, and sperm lengths across butterflies. *Proc. R. Soc. Lond. B* **258**, 247–254 (1994).
19. LaMunyon, C. W. & Ward, S. Evolution of larger sperm in response to experimentally increased sperm competition in *Caenorhabditis elegans*. *Proc. R. Soc. Lond. B* **269**, 1125–1128 (2002).
20. Pitnick, S., Markow, T. A. & Spicer, G. S. Delayed male maturity is a cost of producing large sperm in *Drosophila*. *Proc. Natl Acad. Sci. USA* **92**, 10614–10618 (1995).
21. Pitnick, S. Investment in testes and the cost of making long sperm in *Drosophila*. *Am. Nat.* **148**, 57–80 (1996).
22. Pitnick, S. & Markow, T. A. Male gametic strategies: sperm size, testes size, and the allocation of ejaculate among successive mates by the sperm-limited fly *Drosophila pacifica* and its relatives. *Am. Nat.* **143**, 785–819 (1994).
23. Joly, D., Bressac, C., Devaux, J. & Lachaise, D. Sperm length diversity in Drosophilidae. *Drosoph. Inf. Serv.* **70**, 104–108 (1991).
24. Pitnick, S., Spicer, G. S. & Markow, T. A. Phylogenetic examination of female incorporation of ejaculates in *Drosophila*. *Evolution* **51**, 833–845 (1997).
25. Pitnick, S., Spicer, G. S. & Markow, T. A. How long is a giant sperm? *Nature* **375**, 109 (1995).
26. Jones, A. G., Arguello, J. R. & Arnold, S. J. Molecular parentage analysis in experimental newt populations: the response of mating system measures to variation in the operational sex ratio. *Am. Nat.* **164**, 444–456 (2004).
27. Karr, T. L. & Pitnick, S. The ins and outs of fertilization. *Nature* **379**, 405–406 (1996).
28. Andersson, M. *Sexual Selection* (Princeton Univ. Press, Princeton, New Jersey, 1994).
29. Markow, T. A. & Sawka, S. Dynamics of mating success in experimental groups of *Drosophila melanogaster* (Diptera: Drosophilidae). *J. Insect Behav.* **5**, 375–383 (1992).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Dallai for contributing the scanning electron micrographs of *Drosophila bifurca* for Fig. 1, W. T. Starmer for statistical advice, B. A. Byrnes for technical assistance, W. J. Etges for directions to the *D. bifurca* collection site in Mexico, and C. Hubbell and SUNY Upstate Medical University for providing access to the gamma irradiator. We are also indebted to S. M. Shuster, M. J. Wade, S. J. Arnold, M. Kirkpatrick, G. A. Parker, R. Lande, R. A. Schmedicke, L. L. Wolf, W. T. Starmer, W. D. Brown, J. A. C. Uy and G. T. Miller for discussion of our data and/or comments on an earlier draft of this manuscript. This work was supported by a National Science Foundation Grant to S.P. and A.B.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.B. (acbjork@syr.edu).

LETTERS

Social interactions among epithelial cells during tracheal branching morphogenesis

Amin S. Ghabrial¹ & Mark A. Krasnow¹

Many organs are composed of tubular networks that arise by branching morphogenesis in which cells bud from an epithelium and organize into a tube^{1–3}. Fibroblast growth factors (FGFs) and other signalling molecules have been shown to guide branch budding and outgrowth^{4–7}, but it is not known how epithelial cells coordinate their movements and morphogenesis. Here we use genetic mosaic analysis in *Drosophila melanogaster* to show that there are two functionally distinct classes of cells in budding tracheal branches: cells at the tip that respond directly to Branchless FGF and lead branch outgrowth, and trailing cells that receive a secondary signal to follow the lead cells and form a tube. These roles are not pre-specified; rather, there is competition between cells such that those with the highest FGF receptor activity take the lead positions, whereas those with less FGF receptor activity assume subsidiary positions and form the branch stalk. Competition appears to involve Notch-mediated lateral inhibition that prevents clusters of cells surrounding each sac⁴. There may also be cooperation between budding cells, because in a mosaic epithelium, cells that cannot respond to the chemoattractant, or respond only poorly, allow other cells in the epithelium to move ahead of them.

The *Drosophila* tracheal system develops from epithelial sacs of about 80 cells from which primary, secondary and terminal branches sprout without cell division or cell death^{8,9}. Primary branch sprouting is induced by Branchless (Bnl) FGF, a chemoattractant secreted by clusters of cells surrounding each sac⁴ (Fig. 1a), which activates Breathless (Btl) FGF receptor (FGFR), a receptor tyrosine kinase expressed on tracheal cells¹⁰. Primary branches contain 3–20 cells that organize into a tube as they migrate out from the sac (Fig. 1b). Bnl also induces the expression of secondary branching genes, such as the transcription factor *painted* (*pnt*)¹¹, and specifies terminal cells at the ends of outgrowing branches^{4,8}. Terminal cells ramify in the larva in response to Bnl to form fine terminal branches¹². Other cells at the ends of primary branches become fusion cells that connect with neighbouring branches to form a continuous tracheal network. Terminal and fusion cell fate decisions are also influenced by the Notch, Dpp and Wingless signalling pathways^{13–17}. Dorsal branches, the primary branches that we focus on here, typically consist of five or six cells: two cells near the branch tip, one (DB1) that becomes a terminal cell and another (DB2) that becomes a fusion cell, and three or four cells (DB3–DB6) that form the branch stalk (Fig. 1a, b).

In a genetic mosaic screen (A.S.G., B. P. Levi and M.A.K., unpublished observations), six mutants (724, 788, 1118, 1187, 1476 and 1684) were identified with a subtle phenotype: mosaic branches (+/+ , +/- , -/- cells) were grossly normal, yet homozygous mutant clones (-/- cells) rarely if ever included terminal cells (Fig. 1c–e, and Supplementary Tables S1–S4; data not shown for other primary branches). These were neither general nor terminal-cell-specific lethal mutations because homozygous mutant cells were readily recovered in all other tracheal positions, and there was no decrease in the overall number of cells in mosaic dorsal branches

(5.3 ± 1.1 (mean $\pm$ s.d.) versus 4.9 ± 1.1 in contralateral control branches; $n = 22$ pairs of branches) or the number of terminal cells (98% of both wild-type ($n = 127$) and 724 or 788 mosaic dorsal branches ($n = 291$) had a terminal cell). It was difficult to imagine how mutations could block clone generation in specific cells. It seemed more likely that the mutations caused cells otherwise destined to become terminal cells to switch fates with other tracheal cells.

The six mutations compose a single lethal complementation group that mapped to the left arm of chromosome 3 and failed to complement *breathless*^{LG18}. DNA sequencing identified a single nucleotide change in each mutant resulting in a nonsense or missense mutation in *btl* (Fig. 2a). Five mutations (724, 788, 1118, 1476 and 1684) appear to be null *btl* mutations, whereas the sixth mutation (1187) causes partial loss of function (Fig. 2b and data not shown). Thus, the 'no mutant terminal cells' gene is *btl*.

We quantified the distribution of cells homozygous for *btl* null mutations (724 and 788), or homozygous for a wild-type *btl* allele as a control, in mosaic dorsal branches (Supplementary Tables S1–S4). Control clones were evenly distributed throughout the branch at the expected frequencies; for example, the ratio of stalk-cell to terminal-cell clones was about 3:1. By contrast, *btl*^{-/-} cells showed a nearly complete bias against the DB1 position: the ratio of stalk-cell to terminal-cell clones was 51:1. The three exceptional mutant terminal cells may be cases in which the clone was induced after *btl* began to be expressed, allowing wild-type *btl* gene products to perdure in mutant cells. We recovered hundreds of mosaic branches with one or more *btl*^{-/-} cells present in positions DB2–DB6 without affecting cell or branch morphology. Indeed, branches composed largely or exclusively of *btl*^{-/-} cells, except for a wild-type terminal cell, were morphologically indistinguishable from wild-type branches (Fig. 2c, and Supplementary Tables S3 and S4). Thus, although all tracheal cells normally express *btl*, and the receptor is activated by Bnl in most or all cells of budding branches^{8,18}, the receptor appears to be required in just a single leading cell (DB1). All other cells can migrate normally and form tubes in the absence of *btl*. We conclude that there are two functionally distinct classes of cells in budding primary branches: lead cells, which require Btl FGFR and directly respond to Bnl FGF, and trailing cells, which do not require Btl but follow the lead cell and form the stalk.

What does it take to become the leader? The lead cell (DB1) is specified to become a terminal cell by Bnl–Btl signalling^{4,8}. If terminal cell specification is required, then null mutations in the downstream gene *pnt*, which abolish this function⁸, should have the same effect as *btl* mutations. Cell clones homozygous for *pnt*^{Δ88} or two new *pnt* alleles isolated in our screen (198 and 1318) failed to develop as terminal cells, as expected. However, unlike *btl* mosaic branches, *pnt* mosaic branches often lacked a terminal cell (26% of mosaic branches; $n = 97$). When a terminal cell was missing, there was usually (about 80% of the time) a *pnt*^{-/-} cell in the stalk position nearest the tip (Fig. 2d, and Supplementary Table S5), presumably

¹Howard Hughes Medical Institute, Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305-5307, USA.

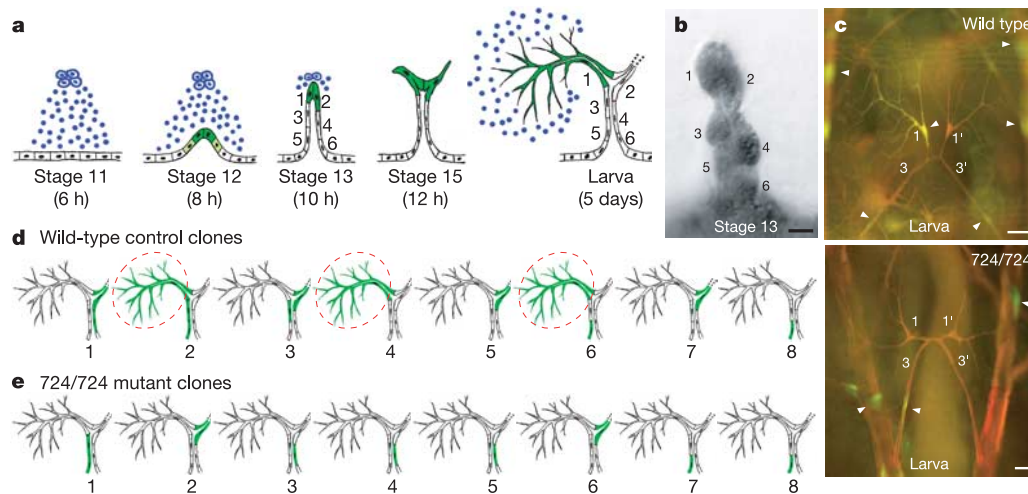


Figure 1 | The 'no mutant terminal cells' phenotype. **a**, Diagram of dorsal branch (DB) budding from tracheal epithelium (black; DB cells numbered 1–6) at the developmental stages and times indicated. Nearby cells (blue) secrete Branchless FGF (blue dots), which activates Breathless (Btl) FGFR on tracheal cells, inducing migration and tube formation. Bnl also induces secondary branching genes (for example *pointed*) in cells (green) that form unicellular secondary branches (stage 15). Subsequently, DB1 (terminal cell) forms terminal branches in response to Bnl expressed by hypoxic larval cells. DB2 (fusion cell) forms a branch that fuses (dotted lines) to a contralateral DB (not shown). DB3–6 cells form DB stalk. **b**, Micrograph of budding DB (stage 13). Nuclei are black; cytoplasm is grey. Cells here are arranged side by side, but subsequently the stalk cells intercalate. Reprinted

with permission (ref. 8). Scale bar, 2.5 μ m. **c**, Fluorescence micrographs of GFP-marked clones in pairs of larval DBs (dorsal view, anterior up). Tracheal cells are labelled with DsRED (red), and homozygous clones are labelled with GFP (green). Arrowheads indicate select clones in DBs and dorsal trunks (large branches flanking DBs). Top panel: control wild-type clones. Bottom panel: homozygous 724 mutant clones. DB1 (1, 1') and DB3 (3, 3') cells in left and right DBs are indicated. Scale bars, 50 μ m.

d, Distribution of control clones (green) in eight representative DBs. Three contain DB1/terminal cell clones (circled). For simplicity, only DBs with five cells (DB1–DB5) are shown. **e**, Distribution of homozygous 724 mutant clones in eight representative DBs. Clones had no morphological defects but were excluded from DB1 positions.

the DB1 cell that failed to differentiate into a terminal cell. This suggests that *pnt*^{-/-} cells are able to assume the lead position but fail to differentiate as terminal cells, and that the bias against *btl* mutant terminal cells is due to the earlier, *pnt*-independent, function of Btl in primary branch budding and outgrowth. If cells lacking Btl cannot migrate in response to Bnl during budding^{8,10,19}, they should not be able to move to the lead position necessary to be selected as a terminal cell. Consistent with this, genetic mosaic analysis of *stumps* (*dofheartbroken*), which encodes a Btl adaptor required for cell migration^{20,21}, showed a dearth of terminal cell clones similar to *btl* (data not shown).

Two results demonstrate that the ability to sense Bnl and migrate in response to it is not enough to become the leader: cells compete for the lead position. The first involves *btl*^{BN} (E796K mutation in the kinase domain), a weak *btl* allele isolated in a separate screen. Unlike *btl*^{-/-} animals, which die in first larval instar and lack virtually all branches¹⁰ (Fig. 2b), *btl*^{BN} homozygotes survived until L3 larval stage or beyond and had a normally patterned tracheal system with a full complement of terminal cells. The only defects detected were a reduced number and altered morphology of terminal branches, presumably due to the dosage-sensitive function of *btl* in terminal branch outgrowth¹² (Fig. 3a–c). The late and subtle phenotype demonstrates that Btl^{BN} protein retains sufficient activity for early migration and terminal cell specification events. However, we found that in genetic mosaic animals, in which *btl*^{BN/BN} cells must compete with *btl*^{BN/+} and *btl*^{+/+} tracheal cells, *btl*^{BN/BN} cells rarely acquired the lead position (DB1) and developed as terminal cells. Indeed, homozygosity for *btl*^{BN} conferred nearly as complete a bias against becoming a terminal cell as total loss of *btl* (Fig. 3d–f, and Supplementary Tables S1 and S6). Thus, Btl activity above the threshold necessary for migration and terminal cell specification is not sufficient to acquire the lead position and become a terminal cell: a cell must have more Btl activity than other cells in the branch.

Similar conclusions derive from a second experiment in which we analysed marked wild-type (*btl*^{+/+}) cells in heterozygous (*btl*^{+/+}) animals. Whereas *btl*^{+/+} clones in control (*btl*^{+/+}) animals were distributed evenly throughout the branch (Fig. 3d), *btl*^{+/+} clones in

btl^{788/+} heterozygotes preferentially localized to the tip (Fig. 3g, and Supplementary Tables S1, S2 and S7). Cells that did not occupy the lead (DB1) position took positions close to the tip (Fig. 3g, and Supplementary Table S7). Similar results were obtained for *btl*^{+/+} clones in animals heterozygous for *btl*¹¹⁸⁷, a partial-loss-of-function allele (Fig. 3h, and Supplementary Tables S1 and S8). Clones mutant for *sprouty*, an FGF feedback inhibitor^{22,23}, also preferentially populated the tip (Supplementary Table S9). Together, the data show that there is competition for the lead position: cells with highest *btl* activity assume positions at or near the tip of the branch, whereas those with less or no activity segregate towards its base.

Because small differences in *btl* dosage or activity affect a cell's ability to compete for the lead, we investigated whether lateral inhibitory mechanisms that amplify small differences in signalling might be operative. Data suggest that the Notch pathway, a lateral signalling pathway implicated in cell specification events including cell fate determination at tracheal branch tips, also affects cell arrangement^{15–17}. *N^{ts}* embryos shifted to the restrictive temperature during budding formed branches in which most DB cells behaved like lead cells, resulting in large clusters of cells congregated at the lead position¹⁷, whereas expression of constitutively active N^{ACT} throughout the tracheal epithelium had the opposite effect, arresting outgrowth and stalling cells near the base of the branch (Fig. 4a–c). We propose that Notch-mediated lateral inhibition among tracheal cells prevents extra cells from assuming the lead position.

Our results provide evidence for social stratification and dynamic social interactions between epithelial cells during branching morphogenesis (Fig. 4d). First, the results show that budding cells are functionally specialized. A cell at the branch tip requires *btl* and leads outgrowth towards the Bnl signalling centre. Trailing cells do not require *btl* but nevertheless follow the lead cell towards the Bnl source. Because tracheal cells do not migrate or form tubes in *btl*^{-/-} animals, trailing cells must receive a secondary signal generated by the lead cell that induces them to migrate and also activates their tubulogenesis programme. This could be a secreted molecule or physical stimulus such as pulling or stretching the trailing cells.

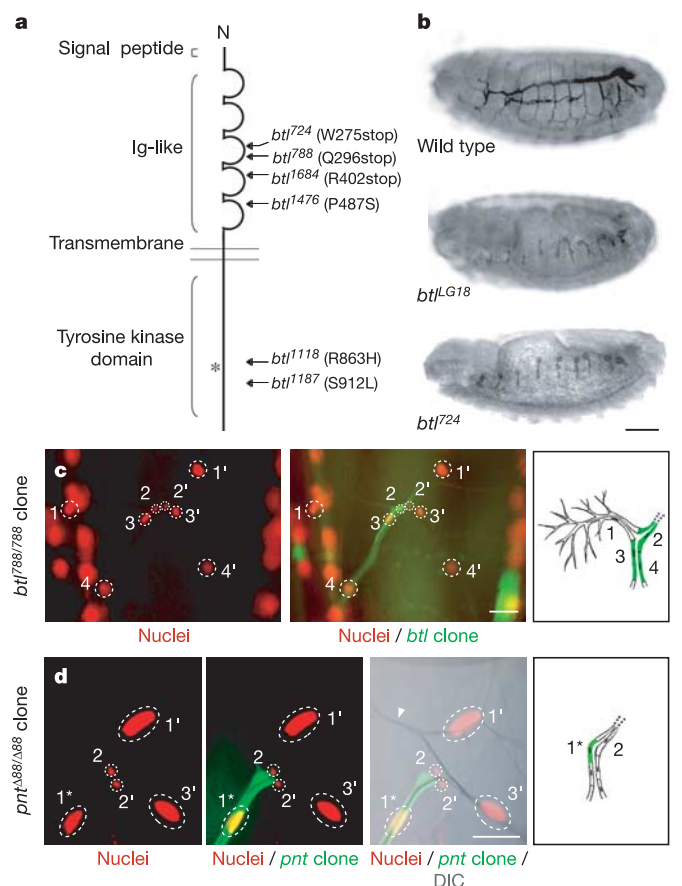


Figure 2 | 'No mutant terminal cells' mutations are loss-of-function mutations in *btl*. **a**, Changes in *btl* coding sequence in 'no mutant terminal cells' mutants. Asterisk, kinase active site. **b**, Wild-type, homozygous *btl*^{LG18} and homozygous *btl*⁷²⁴ embryos (10–13 h old) immunostained with tracheal-specific monoclonal antibody 2A12. Lateral views, anterior left, dorsal up. Note the absence of branches in both mutants. Scale bar, 50 μ m. **c**, Fluorescence micrographs of mosaic DB with large *btl*^{788/788} clone (dorsal view). Tracheal nuclei are labelled with DsRed2nls (red; left and middle panels), and DB cell nuclei are circled (DB1–DB4, mosaic DB; DB1'–DB4', contralateral control DB). (DB1 staining is partly obscured by underlying dorsal trunk staining.) Homozygous *btl*^{788/788} cells are labelled with GFP (green; middle panel) and comprise the entire mosaic DB except DB1 (see diagram, right panel), yet DB has a normal morphology, like the contralateral control DB. Scale bar, 50 μ m. **d**, Mosaic DB containing a null *pointed* ^{Δ 88} clone (dorsal view). First panel: DsRed2nls-labelled nuclei (red) are circled (DB1* and DB2, mosaic DB; DB1'–DB3', contralateral control DB). Second panel: *pnt* ^{Δ 88/ Δ 88} clone (green, DB1*) did not develop as a terminal cell or assume its normal position; instead it formed the distal cell of the stalk. Third panel: red and green channels superimposed on a differential interference contrast image, showing air-filled lumens of DBs. DB1' extends terminal branches (arrowhead) into region normally supplied by missing terminal cell. Fourth panel: diagram of mosaic DB, with *pnt* ^{Δ 88/ Δ 88} clone shaded green. Scale bar, 50 μ m.

Second, these roles are not pre-specified. Rather, there is competition between cells such that those with high Btl FGFR activity become lead cells whereas those with less or no *btl* FGFR activity become trailing cells and form the branch stalk. Competition appears to involve Notch-mediated lateral inhibitory signalling between tracheal cells, and it may also be influenced by positive feedback mechanisms such as increased activation and expression of Btl as cells approach the Bnl source²⁴. Third, there may be cooperation between cells, because in a genetically mosaic epithelium, tracheal cells with less Btl activity allow those with more activity to move ahead of them.

There may be similar social interactions between budding cells in other branching organs. Studies of other branching processes have identified genes selectively expressed in tip cells of budding branches,

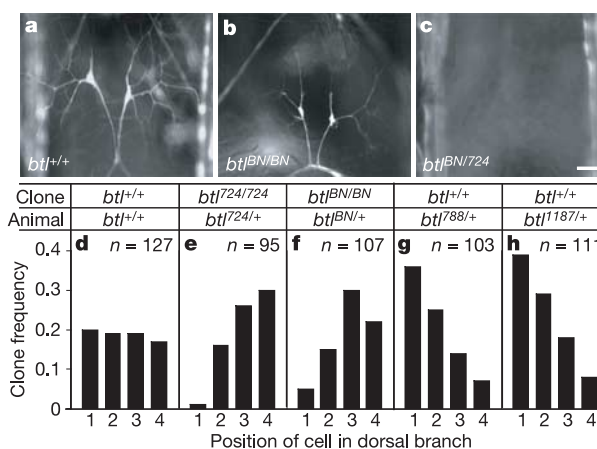


Figure 3 | Effect of *btl* dosage on DB cell position. **a–c**, Fluorescence micrographs of DB pair (dorsal view) in control *btl*^{+/+} larva (**a**), larva homozygous for the weak hypomorphic allele *btl*^{BN} (**b**) and larva transheterozygous for *btl*^{BN} and null *btl*⁷²⁴ allele (**c**). Tracheal cells were labelled by expression of a UAS-GFP transgene driven by *btl*-Gal4 (*btl*-Gal4 > GFP). The *btl*^{BN/BN} larva has normal DB primary and secondary branches, and a normal number of terminal cells, although terminal branching is reduced. *btl*^{BN/724} lacks DBs. Scale bar, 50 μ m. **d–h**, Relative frequency of GFP-marked clones at different DB positions. Genotypes of clones and animals in which clones were induced are indicated. *n*, number of DBs analysed. Because the average number of cells per clone varied in different experiments (**d**, 1.67; **e**, 1.33; **f**, 1.36; **g**, 1.52; **h**, 1.55), clone frequencies shown are relative values normalized to one GFP-marked cell per branch. **d**, Control clones were evenly distributed between positions DB1–DB4. **e**, Clones with no *btl* activity (*btl*^{724/724}) induced in *btl*^{724/+} animals showed a graded distribution, rarely in DB1 position and increasing in frequency towards the branch base. **f**, Clones with mildly reduced *btl* activity (*btl*^{BN/BN}) induced in *btl*^{BN/+} animals showed a similar distribution, even though *btl*^{BN/BN} cells populate all positions in *btl*^{BN/BN} animals (**b**). **g, h**, *btl*^{+/+} clones induced in animals heterozygous for null *btl*⁷⁸⁸ allele (**g**) or moderate *btl*¹¹⁸⁷ allele (**h**) showed the opposite distribution, preferentially in DB1 positions and decreasing in frequency towards the branch base.

and in some cases these cells display morphological specializations indicating that they might actively lead outgrowth²⁵. However, because most budding branches contain hundreds or thousands of cells, it is difficult to track and manipulate individual cells to investigate social behaviours like those described here. Recent analyses of chimaeric *Ret*^{+/Ret} mouse renal ureteric buds in culture²⁶ and *btl* mosaic air sacs²⁷ reveal that cells lacking these receptor tyrosine kinases are excluded from branch tips, indicating that RTK-dependent interactions similar to those described here might be operative in more complex branching events.

METHODS

Fly stocks. *btl*^{LG18}, *spz* ^{Δ 5}, *pnt* ^{Δ 88} and *N^{11N-151}* (abbreviated *N¹⁵*) are described in Flybase (<http://flybase.bio.indiana.edu/>). New *btl* and *pnt* alleles were generated by ethane methyl sulphamate mutagenesis of a third chromosome containing FLP recombinase sites FRT2A and FRT82B (Flybase) and isolated in screens described elsewhere (A.S.G., B. P. Levi and M.A.K., unpublished observations). The Gal4-UAS system²⁸ with *btl*-GAL4 driver (Flybase) was used to express proteins throughout the tracheal system. UAS responders were as follows: UAS-DsRED (A.S.G., B. P. Levi and M.A.K., unpublished observations), which expresses a cytoplasmic red fluorescent protein; UAS-DsRED2nls (M. Galko, G. Fish and M.A.K., unpublished observations), which expresses a nuclear DsRed2 protein; UAS-GFP (Flybase); UASi-GFP (A.S.G., B. P. Levi and M.A.K., unpublished observations), which expresses a double-stranded hairpin RNA that inhibits GFP expression by RNA-mediated interference; and UAS-N^{ACT} (Flybase).

Clone generation and labelling. Heterozygous embryos (1–3 h old) carrying the *hsFLP*¹⁻²² recombinase transgene (Flybase) were reared at 25 °C and then placed at 38 °C for 45–60 min to induce mitotic recombination at FRT sites centromere-proximal to the mutation, generating clones of homozygous mutant cells²⁹. Tracheal cells were marked with a *btl*-GAL4 transgene to drive expression of UAS-DsRED (or UAS-DsRED2nls) and UAS-GFP transgenes. Expression of the latter was blocked by a UASi-GFP transgene present *in trans* to the mutation

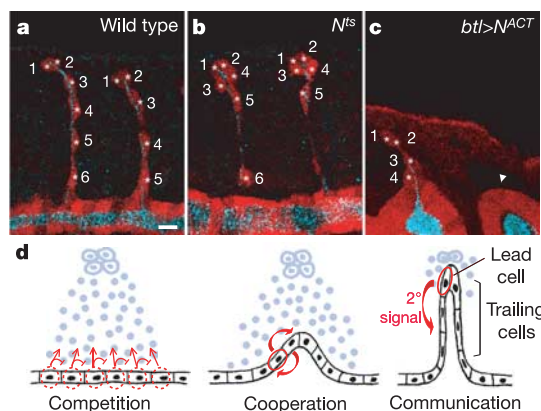


Figure 4 | Effects of Notch (*N*) activity on cell position. **a–c**, Fluorescent micrographs of two DBs (lateral view) in stage 15 wild-type embryo (**a**), *N^{ts}* embryo shifted to non-permissive temperature for 6 h during branch budding (**b**) and *btl-Gal4 > UAS-N^{ACT}* embryo that expressed activated *N* throughout the tracheal system (**c**). All embryos carried *btl-Gal4* and UAS-GFP transgenes and were double-stained for GFP (red; tracheal cell marker) and Vermiform (cyan; luminal marker). **a**, Cells in wild-type DBs are evenly distributed (nuclei are numbered and indicated by asterisks). **b**, *N* inactivation caused the migration of extra cells to the DB tip. **c**, Constitutive *N* activity inhibited outgrowth, particularly in posterior metameres in which some DBs completely failed to bud (arrowhead). Scale bar, 5 μ m. **d**, Social interactions between tracheal cells during budding. The three panels show budding tracheal cells expressing the Btl FGFR moving towards a Bnl FGF signalling centre, as in Fig. 1a. The first panel illustrates cell competition: cells move towards the lead position and inhibit their neighbours from doing the same. The second panel illustrates cell cooperation: a cell with less Btl activity allows one with more to move ahead of it. The third panel illustrates cell communication: the lead cell sends a secondary (2°) signal to the trailing cells, inducing them to follow the lead cell and activating a tubulogenesis programme. Cells also communicate via Notch-mediated signalling as they compete for the lead position (inhibition arrow in first panel).

of interest. Homozygous mutant tracheal cells lack the UASi-GFP transgene and so express GFP as well as DsRED. After clone induction, embryos were returned to 25°C, and after four or five days the third instar larvae were heat-killed (70°C for 3–5 s), mounted in 50% glycerol and observed by fluorescence microscopy to score clone distribution. See Supplementary Methods for complete descriptions of genotypes and clone analysis.

Tracheal cell counts. Mosaic dorsal branches containing at least one homozygous, GFP-expressing cell were scored for presence of terminal (DB1) cells by morphological criteria. Cell numbers in mosaic branches and in contralateral control branches were determined by counting DsRED2nls-labelled nuclei.

Notch experiments. *N^{ts}*; *btl-Gal4*, UAS-GFP/+ embryos were collected at 18°C for 3 h, aged for 10 h, shifted to the non-permissive temperature (30.5°C) for 6 h, then fixed, immunostained and analysed by confocal microscopy. For expression of *N^{ACT}* throughout the tracheal system, *btl-Gal4*, UAS-GFP/+; *btl-Gal4*, UAS-CD8GFP/UAS-*N^{ACT}* embryos were raised at 30.5°C.

Immunostaining and microscopy. Embryo fixation and tracheal staining with monoclonal antibody 2A12 was performed as described⁸. For Fig. 4, embryos were fixed and double-stained with goat anti-GFP (1:200 dilution; Abcam) and anti-Verm³⁰ (1:600 dilution), and revealed by indirect immunofluorescence with fluorescein isothiocyanate-conjugated and CY3-conjugated secondary antibodies. Stained embryos and labelled tracheal clones were analysed and photographed with compound or confocal fluorescence microscopes. Where necessary, a montage of optical sections was assembled (Figs 1c and 4a–c).

Received 30 October 2005; accepted 19 April 2006.

- Hogan, B. L. & Kolodziej, P. A. Organogenesis: molecular mechanisms of tubulogenesis. *Nature Rev. Genet.* **3**, 513–523 (2002).
- Metzger, R. J. & Krasnow, M. A. Genetic control of branching morphogenesis. *Science* **284**, 1635–1639 (1999).
- Affolter, M. *et al.* Tube or not tube: remodeling epithelial tissues by branching morphogenesis. *Dev. Cell* **4**, 11–18 (2003).
- Sutherland, D., Samakovlis, C. & Krasnow, M. A. branchless encodes a *Drosophila* FGF homolog that controls tracheal cell migration and the pattern of branching. *Cell* **87**, 1091–1101 (1996).
- Park, W. Y., Miranda, B., Lebeche, D., Hashimoto, G. & Cardoso, W. V. FGF-10

is a chemotactic factor for distal epithelial buds during lung development. *Dev. Biol.* **201**, 125–134 (1998).

- Weaver, M., Dunn, N. R. & Hogan, B. L. Bmp4 and Fgf10 play opposing roles during lung bud morphogenesis. *Development* **127**, 2695–2704 (2000).
- Pepicelli, C. V., Kispert, A., Rowitch, D. H. & McMahon, A. P. GDNF induces branching and increased cell proliferation in the ureter of the mouse. *Dev. Biol.* **192**, 193–198 (1997).
- Samakovlis, C. *et al.* Development of the *Drosophila* tracheal system occurs by a series of morphologically distinct but genetically coupled branching events. *Development* **122**, 1395–1407 (1996).
- Ghabrial, A., Luschnig, S., Metzstein, M. M. & Krasnow, M. A. Branching morphogenesis of the *Drosophila* tracheal system. *Annu. Rev. Cell Dev. Biol.* **19**, 623–647 (2003).
- Klamt, C., Glazer, L. & Shilo, B. Z. *breathless*, a *Drosophila* FGF receptor homolog, is essential for migration of tracheal and specific midline glial cells. *Genes Dev.* **6**, 1668–1678 (1992).
- Klamt, C. The *Drosophila* gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development* **117**, 163–176 (1993).
- Jarecki, J., Johnson, E. & Krasnow, M. A. Oxygen regulation of airway branching in *Drosophila* is mediated by branchless FGF. *Cell* **99**, 211–220 (1999).
- Llimargas, M. Wingless and its signalling pathway have common and separable functions during tracheal development. *Development* **127**, 4407–4417 (2000).
- Chihara, T. & Hayashi, S. Control of tracheal tubulogenesis by Wingless signaling. *Development* **127**, 4433–4442 (2000).
- Llimargas, M. The Notch pathway helps to pattern the tips of the *Drosophila* tracheal branches by selecting cell fates. *Development* **126**, 2355–2364 (1999).
- Ikeya, T. & Hayashi, S. Interplay of Notch and FGF signaling restricts cell fate and MAPK activation in the *Drosophila* trachea. *Development* **126**, 4455–4463 (1999).
- Steneberg, P., Hemphala, J. & Samakovlis, C. Dpp and Notch specify the fusion cell fate in the dorsal branches of the *Drosophila* trachea. *Mech. Dev.* **87**, 153–163 (1999).
- Gabay, L., Seger, R. & Shilo, B. Z. MAP kinase in situ activation atlas during *Drosophila* embryogenesis. *Development* **124**, 3535–3541 (1997).
- Ribeiro, C., Ebner, A. & Affolter, M. *In vivo* imaging reveals different cellular functions for FGF and Dpp signaling in tracheal branching morphogenesis. *Dev. Cell* **2**, 677–683 (2002).
- Vincent, S., Wilson, R., Coelho, C., Affolter, M. & Leptin, M. The *Drosophila* protein Dof is specifically required for FGF signaling. *Mol. Cell* **2**, 515–525 (1998).
- Imam, F., Sutherland, D., Huang, W. & Krasnow, M. A. stumps, a *Drosophila* gene required for fibroblast growth factor (FGF)-directed migrations of tracheal and mesodermal cells. *Genetics* **152**, 307–318 (1999).
- Hacohen, N., Kramer, S., Sutherland, D., Hiromi, Y. & Krasnow, M. A. sprouty encodes a novel antagonist of FGF signaling that patterns apical branching of the *Drosophila* airways. *Cell* **92**, 253–263 (1998).
- Casci, T., Vinos, J. & Freeman, M. Sprouty, an intracellular inhibitor of Ras signaling. *Cell* **96**, 655–665 (1999).
- Ohshiro, T., Emori, Y. & Saigo, K. Ligand-dependent activation of *breathless* FGF receptor gene in *Drosophila* developing trachea. *Mech. Dev.* **114**, 3–11 (2002).
- Gerhardt, H. G., Alitalo, K., Shima, D., Betsholtz, C. *et al.* VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia. *J. Cell Biol.* **161**, 1163–1177 (2003).
- Shakya, R., Watanabe, T. & Costantini, F. The role of GDNF/Ret signaling in ureteric bud cell fate and branching morphogenesis. *Dev. Cell* **8**, 65–74 (2005).
- Cabernard, C. & Affolter, M. Distinct roles for two receptor tyrosine kinases in epithelial branching morphogenesis in *Drosophila*. *Dev. Cell* **9**, 831–842 (2005).
- Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Golic, K. G. Site-specific recombination between homologous chromosomes in *Drosophila*. *Science* **252**, 958–961 (1991).
- Luschnig, S., Batz, T., Armbruster, K. & Krasnow, M. A. serpentine and vermiform encode matrix proteins with chitin binding and deacetylation domains that limit tracheal tube length in *Drosophila*. *Curr. Biol.* **16**, 186–194 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank B. Levi for help with isolation of ‘no mutant terminal cell’ mutants and for discussion, and S. Artavanis, M. Galko, S. Luschnig and S. Toering for fly stocks and antisera. This work was supported by an NIH NRSA fellowship (A.S.G.) and a grant from the N.I.H. M.A.K. is an investigator of the Howard Hughes Medical Institute.

Author Contributions A.S.G. conceived, designed and performed experiments and analysed data. M.A.K. advised on the above. A.S.G. and M.A.K. wrote the paper.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.A.K. (krasnow@cmgm.stanford.edu)

LETTERS

Importance of SoxE in neural crest development and the evolution of the pharynx

David W. McCauley^{1†} & Marianne Bronner-Fraser¹

The neural crest, a defining character of vertebrates¹, is of prime importance to their evolutionary origin². To understand neural crest evolution, we explored molecular mechanisms underlying craniofacial development in the basal jawless vertebrate, sea lamprey (*Petromyzon marinus*), focusing on the *SoxE* (*Sox8*, *Sox9* and *Sox10*) gene family. In jawed vertebrates, these are important transcriptional regulators of the neural crest³, and the loss of *Sox9* causes abnormal craniofacial development^{4,5}. Here we report that two lamprey *SoxE* genes are expressed in migrating neural crest and crest-derived prechondrocytes in posterior branchial arches, whereas a third paralogue is expressed later in the perichondrium and mandibular arch. Morpholino knock-down of *SoxE1* reveals that it is essential for posterior branchial arch development, although the mandibular arch is unaffected. The results show that chondrogenic function of *SoxE* regulators can be traced to the lamprey–gnathostome common ancestor and indicate that lamprey *SoxE* genes might have undergone independent duplication to have distinct functions in mandibular versus caudal branchial arches. This work sheds light on the homology of vertebrate branchial arches and supports their common origin at the base of vertebrates.

Neural crest cells, a hallmark of the vertebrate 'new head'¹, were critical to the evolution of the vertebrate body plan. As the basalmost extant vertebrate, lampreys can provide important insights into evolution of neural crest developmental mechanisms. Lamprey embryos are accessible for developmental studies and have neural crest contributions to their branchial arches^{6–9}. Long-noted differences between lamprey and gnathostome branchial arches have raised questions about their homology^{10–13}. To address this question we investigated mechanisms underlying chondrogenesis and branchial arch formation in lamprey. Here we report functional differences between mandibular (first) and posterior arches related to differences in *SoxE* deployment arising after duplication of a single ancestral *SoxE* gene.

We identified three *SoxE* genes in *P. marinus* and used a Clustal¹⁴ amino-acid alignment to determine their orthology or paralogy. *SoxE1* and *SoxE2* were basal to gnathostome *SoxE* genes, with *SoxE1* equidistant from all gnathostome *SoxE* genes and *SoxE3* showing sequence homology to *Sox9* (Supplementary Information).

In situ hybridization¹⁵ at progressive stages from neurulation to formation of the pharyngeal arches revealed expression patterns of *SoxE* genes during neural crest development. *SoxE1* and *SoxE2* were both expressed in neural folds and later in migrating neural crest (Supplementary Information), including those of posterior branchial arches (Fig. 1a, d), but were absent from the mandibular arch. *SoxE3* expression, absent early, was observed later in mesenchyme of all arches including the velum of the mandibular arch (Fig. 1g). After arch formation, all three *SoxE* genes were expressed in the rostralateral portion of the arches that form branchial basket

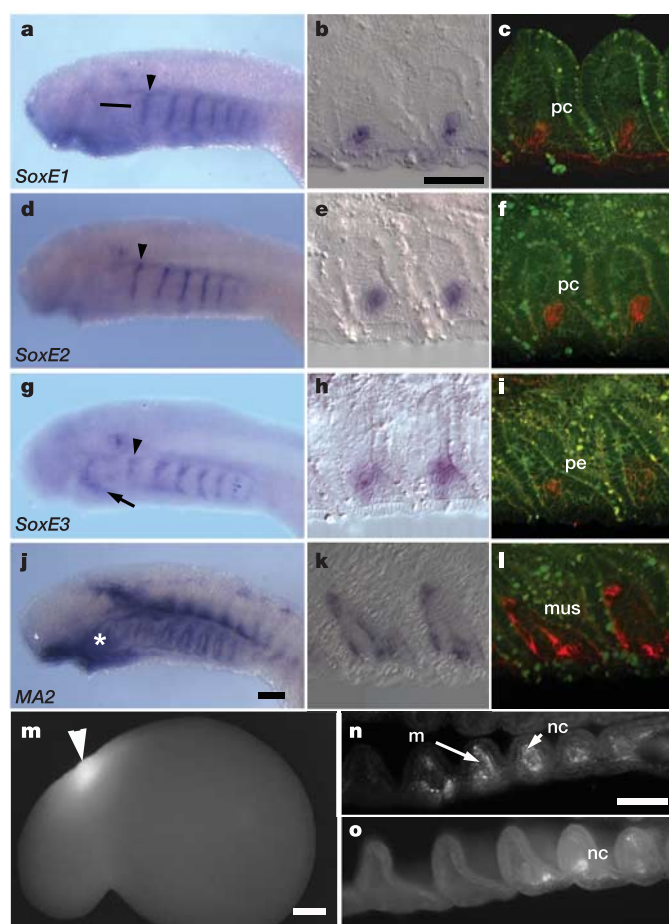


Figure 1 | SoxE RNA transcripts in neural crest within branchial arches. **a–l**, Whole mounts *in situ* (**a**, **d**, **g**, **j**) stained with NBT/BCIP, sectioned and imaged with differential interference contrast (**b**, **e**, **h**, **k**) and confocal microscopy (**c**, **f**, **i**, **l**). **a–c**, *SoxE1* expression at stage 25 in chondrocytes within each arch posterior to the mandibular. The line in **a** indicates the approximate plane of section in **b** and **c**. **d–f**, *SoxE2* expression in chondrocytes at stage 25. **g–i**, *SoxE3* expression in mandibular arch (arrow) and in perichondrial cells within posterior arches. **j–l**, MA2 expression in velar and branchial muscles surrounding mucocartilage (asterisk in mandibular arch) and branchial arch cartilage (posterior arches). **m**, Stage-21 embryo labelled with DiI at the hindbrain level during neural crest migration. **n**, DiI-labelled cells surround a mesodermal core in a stage-24 embryo. **o**, DiI-labelled neural crest occupies the same position as cells expressing *SoxE1* and *SoxE2* at stage 25 (see **b**, **c**, **e**, **f**). Abbreviations: m, mesoderm; mus, muscle; nc, neural crest; pc, prechondrocyte; pe, perichondrium. Scale bars, 0.1 mm (**a**, **d**, **g**, **j**); 0.05 mm (**b**, **c**, **e**, **f**, **h**, **i**, **k**, **l**); 0.1 mm (**m–o**).

¹Division of Biology, California Institute of Technology, Pasadena, California 91125, USA. [†]Present address: Department of Zoology, University of Oklahoma, Norman, Oklahoma 73019, USA.

cartilage (Fig. 1b, e, h). Muscle-specific actin expression¹⁶ (MA2; Fig. 1j–l) surrounded chondrocytes in posterior arches (Fig. 1k, l) and was present in velar muscle surrounding mucocartilage within the mandibular arch¹⁶ (asterisk in Fig. 1j). To examine expression at single-cell resolution, we used a previously unreported optical property of the alkaline phosphatase substrates nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) to reveal important differences in *SoxE* distributions within branchial arches (Fig. 1c, f, i). Whereas *SoxE1* and *SoxE2* were expressed in prechondrocytes that condense into branchial basket cartilage bars (Fig. 1c, f), *SoxE3* was expressed only in the perichondrium surrounding the chondrocytes (Fig. 1i and Supplementary Movies).

To establish the developmental origin of *SoxE*-expressing cells, stage-21 embryos were labelled with the chloromethylbenzamido derivative of 1,1'-diiodo-3,3',3'-tetramethylindocarbocyanine (CM-DiI) subjacent to the dorsal midline epidermis (Fig. 1m) at the level of the presumptive pharynx. By stage 24, labelled neural crest populated branchial arches, surrounding the central mesoderm core both medially and laterally^{6,9} (Fig. 1n). Subsequently, DiI-labelled cells (Fig. 1o) overlapped *SoxE1* and *SoxE2* expression laterally in the branchial arches (Fig. 1b, e), confirming their neural crest origin. Thus, rostralateral positioning of cartilage bars seems to occur independently of and later than neural crest migration into the arches¹¹.

That *SoxE*-expressing neural crest invades branchial arches similarly in lamprey and other vertebrates raises the possibility that one or more *SoxE* genes might have an ancestral role in neural crest and pharyngeal arch development. To address this, we used two different morpholino antisense oligonucleotides (MOs) to disrupt *SoxE1* translation. By stage 26, knock-down of *SoxE1* protein led to profound defects in the pharyngeal arches, similar to effects reported in *Xenopus*⁴ (Table 1). The length of the pharynx was reduced and arches were lost or greatly reduced in 52% of zygotes injected with MO1 (Fig. 2a, b) and 29% of surviving embryos with MO2. Injection of both MO1 and MO2 resulted in 100% mortality before the completion of neurulation, indicating that *SoxE1* might be required before neural-tube closure. To confirm the specificity of the morpholino, we used a *SoxE1* morpholino with five base-pair mismatches, and two unrelated morpholinos (*Gallus gallus* Pax7 and *P. marinus* Pax2/5/8)¹⁷. Neither mismatch (0/70) nor chick morpholinos gave any phenotype, whereas the Pax2/5/8 morpholino resulted in a phenotype distinct from that of *SoxE1*MO (not shown), confirming its specificity. By targeting a single blastomere at the two-cell stage (Fig. 2c), MO was sequestered in one side of late-neurula-stage embryos (Fig. 2d). Subsequently, branchial arches were lost from the injected side (Fig. 2e, f) in 41% and 31% of embryos receiving MO1 and MO2, respectively.

An antibody against lamprey *SoxE* confirmed *SoxE1* depletion. Immunopositive cells were observed in prechondrocytes, perichondrium and lateral mesenchyme of the branchial arches (Fig. 2f–h), reflecting the combined *in situ* patterns of *SoxE1*, *SoxE2* and *SoxE3* (Fig. 2h). After *SoxE1*MO injection at the two-cell stage, decreased protein expression was found on the injected side (Fig. 2g), with concomitant reduction in or loss of endodermal evagination resulting in the absence of pharyngeal pouches (Fig. 2f, g). However, a lateral

domain of expression reminiscent of *SoxE3* mRNA distribution persisted (Fig. 2g).

Morphological effects of *SoxE1*MO on the development of branchial arches were examined with markers for cartilage and mesoderm, *SoxE2* (Fig. 2i) and *SoxE3* (Fig. 2j), and a marker for muscle-specific actin, MA2 (ref. 16) (Fig. 2k). After *SoxE1*MO injection, *SoxE2* expression was lost from the pharyngeal region (Fig. 2i), indicating a loss of neural-crest-derived chondrocytes, although *SoxE2* persisted more caudally. This observation indicates that *SoxE1* might regulate *SoxE2*. In contrast to *SoxE2*, *SoxE3* expression was unaffected (Fig. 2j). Although endodermal pouches failed to fuse with the ectoderm on the MO-injected side, *SoxE3*

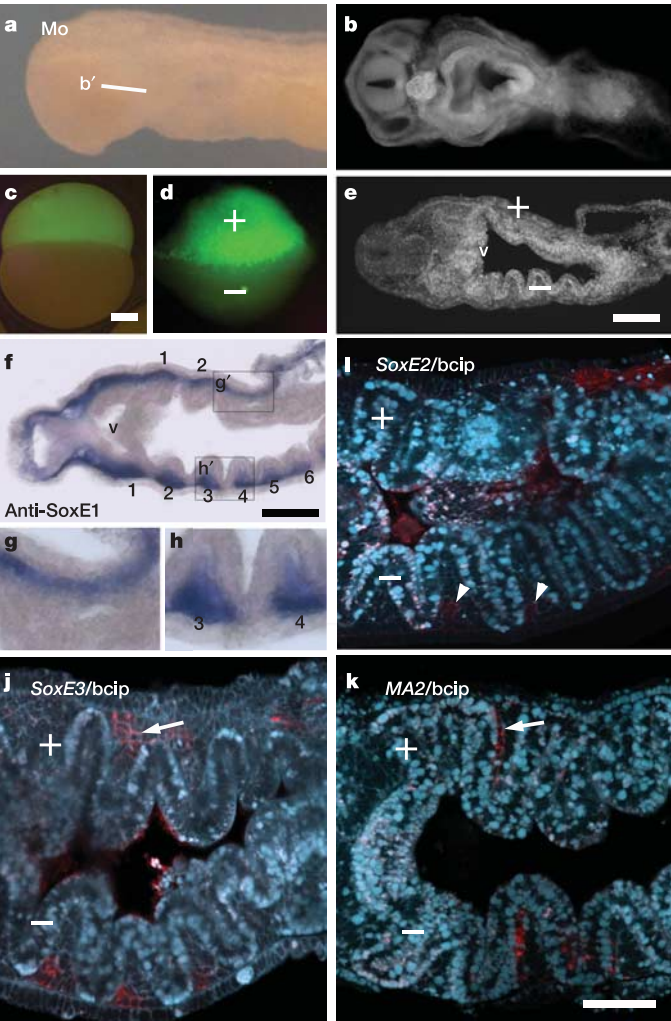


Figure 2 | *SoxE1* MO effects on pharyngeal development. **a–e**, MO injection into the zygote (**a**, **b**) or the single blastomere of a two-cell embryo (**c–e**). **a**, Lateral view, stage 25. **b**, Horizontal section through pharynx in **a**. **c**, Fluorescein isothiocyanate (FITC)-labelled MO injected into a single blastomere at the two-cell stage. **d**, MO in half embryo at stage 21. **e**, Stage 26 embryo lacking branchial arches on the injected side. **f–h**, *SoxE1* immunohistochemistry in a stage-25 embryo after MO injection (as in **c**). Branchial arches are labelled 1 and 2 on the MO-injected side, and 3–6 on the control side. MO-injected half (inset **g'** and **g**) compared with control side (inset **h'** and **h**). **i–k**, NBT/BCIP fluorescence (bcip), showing pharyngeal expression of *SoxE2* (**i**), *SoxE3* (**j**) and MA2 (**k**) after injection of *SoxE1*MO as in **c**. *SoxE2* expression is present on the control side only (arrowheads in **i**); *SoxE3* and MA2 expression are present on the injected side where fully formed arches and pharyngeal pouches are lacking (arrows in **j**, **k**). Key: +, MO-injected side of embryo; –, control side of embryo; v, velum. NBT/BCIP fluorescence is shown in red; lamprey autofluorescence is shown in blue, FITC-labelled *SoxE1*MO is shown in green. Scale bars, 0.1 mm (**a**, **b**, **e**); 0.05 mm (**c**, **d**); 0.1 mm (**f**, **i–k**).

Table 1 | Effects of morpholino injection on branchial arch development

Treatment	Zygote		Embryo	
	Reduced/total	%	Reduced/total	%
MO1	47/90	52	14/34	41
MO2	31/107	29	15/48	31
MO1 + MO2	*	*	*	*
Control MO	0/70	0	–	–

Results show the numbers and percentages of embryos with reduction or loss of branchial arches out of the total that survived to stage 26 after injection of *PmSoxE1* morpholino antisense oligonucleotide into zygotes or into one blastomere of a two-cell-stage embryo. *Lethal.

expression remained, despite the lack of cartilage stacks in this region (Fig. 2j). In some cases the pharyngeal endoderm abutted the ectoderm and the mesodermal core was absent on the morphant side (Fig. 2i, k). Despite the absence of chondrocytes, MA2 was seen in the rostral region of arches adjacent to the endoderm, similarly to the control side (Fig. 2k).

After morpholino knock-down, velar morphology (Fig. 2e, f) and *SoxE3* expression remained unaffected in the first branchial arch, indicating that *SoxE1* might not be involved in mandibular arch formation. Thus, early partitioning of *SoxE* function among paralogues after their duplication at the base of vertebrates might have been crucial for segregating the developmental programme of the first arch from that of the caudal pharyngeal arches that form the branchial basket. Similarly, a subset of lamprey *Dlx* genes (*DlxA*, *DlxC* and *DlxD*) are expressed in premigratory and migratory neural crest, whereas *DlxB* is present in the posterior arches only after neural crest migration into the branchial-arch-forming region⁸.

In vertebrates, gill arches are supported by segmental sets of cartilages located medial to the mesoderm in gnathostomes. In contrast, cartilage bars in lamprey are located laterally. This topography, coupled with *Dlx* expression as a neural crest marker⁸, led Kimmel *et al.*¹¹ to propose an 'outside-in' hypothesis of neural crest migration into the arches such that neural crest migration occurs first laterally and subsequently medially to form a ring around the mesoderm. This medial migration was suggested to be missing in lamprey. In contrast, Damas¹⁸ indicated that cells of ectomesenchymal origin were located medial to the mesodermal core in lamprey. Our labelling study with DiI supports the observations of Damas, that neural crest migration into lamprey branchial arches parallels that observed in gnathostomes⁹ and indicates that development of arches in the early vertebrate might have involved a common neural crest origin and a homologous regulatory mechanism involving *SoxE* genes. Our data support the homology of pharyngeal elements between lamprey and gnathostomes. Thus, differences between the mandibular arch (velum) and posterior arch (branchial basket) derivatives may depend on differential *SoxE* gene deployment, in addition to changes in *Hox* gene expression, as has been suggested^{19,20}. In gnathostomes the mandibular arch contributes to the upper and lower jaws²¹, whereas in lampreys it contributes to the velum and lower lip²². Loss of *SoxE1* function had no effect on the mandibular arch or on *SoxE3* expression. That the development of the mandibular arch occurs independently of caudal arches indicates that duplication of *SoxE* genes early in the vertebrate lineage might have been an important step in decoupling the development of an initially homonomous series of branchial arches in the vertebrate ancestor. This uncoupling gave rise to the homologous branchial arches of modern lampreys and gnathostomes that now develop heteronomously along the anteroposterior axis.

Thus, the expression and function of lamprey *SoxE* genes indicates a possible ancestral role in neural crest and pharyngeal arch development (Supplementary Information). Our results indicate the following: that the common lamprey/gnathostome ancestor possessed a series of branchial arches derived from the neural crest; that *SoxE* genes have independent roles in the morphogenesis of pharyngeal and mandibular arches; and that the neural crest has a crucial function in proper patterning of the pharynx^{23,24}.

METHODS

In situ hybridization and microscopy. RNA expression was determined by whole-mount *in situ* hybridization as described¹⁵. Embryos to be imaged were mounted in 5% low-melt agarose, sectioned at 30–50 μ m with a Vibratome 1000 and mounted in phosphate-buffered saline for viewing. Embryos were imaged with either differential interference contrast optics or fluorescence of DiI with a rhodamine filter set on a Zeiss Axioskop2 microscope. Alternatively, after *in situ* hybridization and sectioning, NBT/BCIP precipitate was detected by confocal microscopy with a Zeiss 510 confocal microscope. Lamprey autofluorescence was excited at 488 nm and emission was detected through a BP500–530 filter. The NBT/BCIP precipitate was excited at 633 nm and the resultant light was detected

through a LP650 filter. Images were imported into Adobe Photoshop for formatting. Complete methods are described in Supplementary Information.

Received 14 December 2005; accepted 3 March 2006.

- Gans, C. & Northcutt, R. G. Neural crest and the origin of vertebrates: a new head. *Science* **220**, 268–274 (1983).
- Holland, N. D. & Chen, J.-Y. Origin and early evolution of the vertebrates: new insights from advances in molecular biology, anatomy, and palaeontology. *BioEssays* **23**, 142–151 (2001).
- Cheung, M. & Briscoe, J. Neural crest development is regulated by the transcription factor *Sox9*. *Development* **130**, 5681–5693 (2003).
- Spokony, R. F., Aoki, Y., Saint-Germain, N., Magner-Fink, E. & Saint-Jeannet, J. P. The transcription factor *Sox9* is required for cranial neural crest development in *Xenopus*. *Development* **129**, 421–432 (2002).
- Yan, Y. L. *et al.* A zebrafish *sox9* gene required for cartilage morphogenesis. *Development* **129**, 5065–5079 (2002).
- Horigome, N. *et al.* Development of cephalic neural crest cells in embryos of *Lampetra japonica*, with special reference to the evolution of the jaw. *Dev. Biol.* **207**, 287–308 (1999).
- Langille, R. M. & Hall, B. K. Role of the neural crest in development of the trabeculae and branchial arches in embryonic sea lamprey, *Petromyzon marinus* (L.). *Development* **102**, 301–310 (1988).
- Neidert, A. H., Virupannavar, V., Hooker, G. W. & Langeland, J. A. Lamprey *Dlx* genes and early vertebrate evolution. *Proc. Natl Acad. Sci. USA* **98**, 1665–1670 (2001).
- McCauley, D. W. & Bronner-Fraser, M. Neural crest contributions to the lamprey head. *Development* **130**, 2317–2327 (2003).
- Janvier, P. *Early Vertebrates* (Oxford Univ. Press, Oxford, 1996).
- Kimmel, C. B., Miller, C. T. & Keynes, R. J. Neural crest patterning and the evolution of the jaw. *J. Anat.* **199**, 105–120 (2001).
- Mallatt, J. Early vertebrate evolution—pharyngeal structure and the origin of gnathostomes. *J. Zool.* **204**, 169–183 (1984).
- Schultze, M. *Die Entwicklungsgeschichte von Petromyzon planeri*, (Haarlem, 1856).
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. Clustal-W—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
- Tomsa, J. M. & Langeland, J. A. *Otx* expression during lamprey embryogenesis provides insights into the evolution of the vertebrate head and jaw. *Dev. Biol.* **207**, 26–37 (1999).
- Kusakabe, R., Takechi, M., Tochinai, S. & Kuratani, S. Lamprey contractile protein genes mark different populations of skeletal muscles during development. *J. Exp. Zool. B* **302**, 121–133 (2004).
- McCauley, D. W. & Bronner-Fraser, M. Conservation of Pax gene expression in ectodermal placodes of the lamprey. *Gene* **287**, 129–139 (2002).
- Damas, H. Observations sur le développement des ganglions craniens chez *Lampetra fluviatilis* (L.). *Arch. Biol. (Liege)* **62**, 65–95 (1951).
- Takio, Y. *et al.* Evolutionary biology: lamprey *Hox* genes and the evolution of jaws. *Nature Brief Communication* arising, doi:10.1038/nature02616 (20 May 2004).
- Coates, M. I. & Cohn, M. J. Fins, limbs and tails: outgrowths and axial patterning in vertebrate evolution. *Bioessays* **20**, 371–381 (1998).
- Cerny, R. *et al.* Developmental origins and evolution of jaws: new interpretation of 'maxillary' and 'mandibular'. *Dev. Biol.* **276**, 225–236 (2004).
- Kuratani, S., Murakami, Y., Nobusada, Y., Kusakabe, R. & Hirano, S. Developmental fate of the mandibular mesoderm in the lamprey, *Lethenteron japonicum*: Comparative morphology and development of the gnathostome jaw with special reference to the nature of the trabecula cranii. *J. Exp. Zool. B Mol. Dev. Evol.* **302**, 458–468 (2004).
- Graham, A. The development and evolution of the pharyngeal arches. *J. Anat.* **199**, 133–141 (2001).
- Veitch, E., Begbie, J., Schilling, T. F., Smith, M. M. & Graham, A. Pharyngeal arch patterning in the absence of neural crest. *Curr. Biol.* **9**, 1481–1484 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Langeland for providing embryonic cDNA libraries; D. Meulemans for *SoxE* oligonucleotide primer sequences, T. Sauka-Spengler for *SoxE2* and *SoxE3* library screens, R. Bergstedt and the staff at Hammond Bay Biological Station for facilities and technical assistance; R. Kusakabe for providing the *LjMA2* DNA construct, and L. Trinh for assistance with confocal microscopy imaging techniques. This work was supported by a grant to M.B.F. from the National Aeronautics and Space Administration.

Author Information *PmSoxE1*, *PmSoxE2*, *PmSoxE3* and *PmSoxF* sequences are deposited in GenBank under accession numbers AY830453, DQ328983, DQ328984 and AY830454, respectively. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare that they have no competing financial interests. Correspondence and requests for materials should be addressed to D.W.M. (dwmccauley@ou.edu).

A dynamic role for the mushroom bodies in promoting sleep in *Drosophila*

Jena L. Pitman¹, Jermaine J. McGill¹, Kevin P. Keegan¹ & Ravi Allada¹

The fruitfly, *Drosophila melanogaster*, exhibits many of the cardinal features of sleep, yet little is known about the neural circuits governing its sleep¹. Here we have performed a screen of GAL4 lines expressing a temperature-sensitive synaptic blocker *shibire*^{ts1} (ref. 2) in a range of discrete neural circuits, and assayed the amount of sleep at different temperatures. We identified three short-sleep lines at the restrictive temperature with shared expression in the mushroom bodies, a neural locus central to learning and memory³. Chemical ablation of the mushroom bodies also resulted in reduced sleep. These studies highlight a central role for the mushroom bodies in sleep regulation.

Evidence indicates that the fruitfly is a valid sleep model¹. Flies exhibit long periods of immobility accompanied by increased arousal thresholds^{4,5}. These states are modulated by drugs known to regulate mammalian sleep^{4,6} and are correlated with brain activity^{7,8}. Fly sleep is homeostatically regulated; that is, flies recover lost sleep, a process known as sleep rebound^{4,5,9}. Initial genetic analyses of fly sleep have implicated a variety of molecular pathways^{5,10–12}. It is not clear whether these pathways regulate sleep through their effects on specialized neural circuits in the brain or indirectly through other loci. Moreover, it is unclear whether sleep is an active process in flies, as in mammals.

To identify brain regions important for sleep in *Drosophila*, we exploited a genetically encoded temperature-sensitive blocker of synaptic transmission, *shibire*^{ts1} (*shi*^{ts1}). *shibire* encodes a *Drosophila* orthologue of dynamin, a GTPase that is essential for synaptic vesicle recycling^{13–15}. At elevated temperatures, *shi*^{ts1} can rapidly block synaptic transmission in neurons^{2,16}. We targeted *shi*^{ts1} expression in defined brain regions with the use of the binary GAL4/UAS system². We selected GAL4 lines from a collection that previously had been shown to drive expression in the adult brain¹⁷, but demonstrating varied expression patterns (Supplementary Table S1). This approach allowed the rapid manipulation of discrete neural ensembles and assessment of the consequences on sleep.

Flies were subjected to temperature cycling of 12 h at 29 °C and 12 h at 21 °C. Of the 92 lines tested, 6 showed decreases in total sleep at 29 °C (Fig. 1a) but not at 21 °C (Fig. 1b). We also found lines with increased sleep at 29 °C (Supplementary Table S1). Given our interest in sleep-promoting circuits, we focused on the short-sleep GAL4 lines (SSL). In temperature cycling, sleep reductions relative to controls are evident soon after temperature shifts to 29 °C (1 h or less), indicating the likelihood of direct effects (Fig. 1c, d). Because reduced sleep may be a consequence of response to temperature shifts, we tested SSL lines maintained at a restrictive (29 °C) temperature. Of the six selected lines, we found three (c253, c747 and c758) with reduced sleep and subsequently focused on them. These phenotypes are temperature dependent because sleep is largely unaffected at 21 °C (Fig. 1f). At 29 °C, sleep is reduced during light and dark periods (Fig. 1g, h) and in both males and females (data not shown). Moreover, these effects are not likely to be due to differences in genetic back-

ground because sibling progeny of heterozygous parents (GAL4/+, UAS*shi*^{ts1}/+; see Supplementary Methods) revealed reduced sleep at 29 °C in GAL4/UAS*shi*^{ts1} flies (Supplementary Fig. S1a, b).

We assessed GAL4 expression patterns by examining driven patterns of green fluorescent protein (GFP) and membrane-linked GFP expression (UAS-mCD8-GFP). Consistent with previous

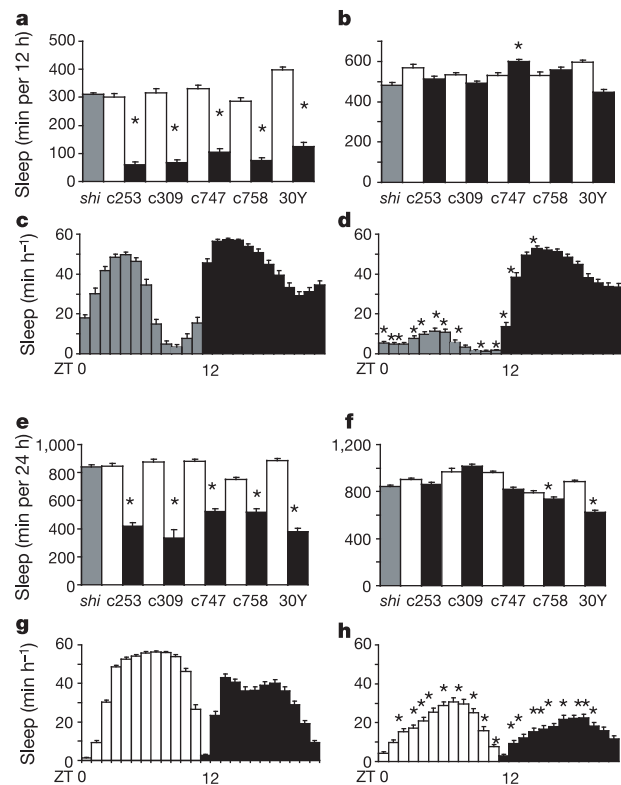


Figure 1 | Sleep in short-sleep GAL4 lines. **a–d**, Temperature cycling. **a**, **b**, Average total sleep at 29 °C (**a**) and 21 °C (**b**) during temperature cycling. Grey bars, +/UAS*shi*^{ts}; white bars, GAL4/+; black bars, GAL4/UAS*shi*^{ts}. **c**, **d**, Hourly sleep for c253/+ (**c**) and c253/UAS*shi*^{ts} flies (**d**). Grey bars, 29 °C (ZT 0–12); black bars, 21 °C (ZT 12–24). **e–h**, Constant temperature. **e**, **f**, Average total sleep during constant 29 °C (**e**) or 21 °C (**f**) conditions. Grey bars, +/UAS*shi*^{ts}; white bars, GAL4/+; black bars, GAL4/UAS*shi*^{ts}. **g**, **h**, Hourly sleep for c253/+ (**g**) and c253/UAS*shi*^{ts} flies (**h**) during constant 29 °C conditions. White bars, light (ZT 0–12); black bars, dark (ZT 12–24). Asterisk, GAL4/UAS*shi*^{ts} combination significantly different from both GAL4/+ and +/UAS*shi*^{ts} (one-way ANOVA, $P < 0.05$). In **a–d**, $n = 33–124$ and the number of experiments was 3–12; in **e–h**, $n = 24–170$ and the number of experiments was two to nine. Error bars indicate s.e.m.

¹Department of Neurobiology and Physiology, Northwestern University, Evanston, Illinois 60208, USA.

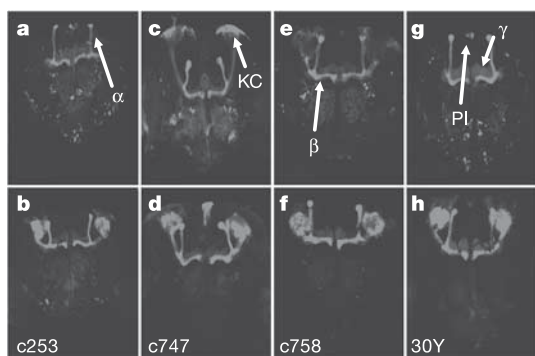


Figure 2 | Short-sleep GAL4 lines share expression within the MB. **a, c, e, g,** UAS-GFP expression. **b, d, f, h,** UAS-mCD8-GFP expression. Expression patterns of short-sleep GAL4 lines c253 (**a, b**), c747 (**c, d**), c758 (**e, f**) and 30Y (**g, h**) are shown. The α , β and γ lobes are labelled. KC, Kenyon cell body layer; PI, pars intercerebralis. Expression patterns were examined in about ten brains per genotype, in two independent experiments.

reports¹⁷, the most prominent areas of shared expression were the mushroom bodies (MB), which are neuropil structures central to some forms of long-term memory¹⁸ (Fig. 2). We then selected other MB-expressing GAL4 lines and found that lines 30Y and c309 GAL4 (ref. 17) showed a robust SSL phenotype in combination with UAS shi^{ts1} (Fig. 1a, b, e, f). The MB GAL4 line 247 (ref. 19) in combination with UAS shi^{ts1} revealed reduced sleep only during the early morning at 29 °C (data not shown). In contrast, line 17D (ref. 20) as well as other MB-expressing lines (Supplementary Table S1) exhibited wild-type or even increased sleep at 29 °C. These lines may differ in their extent of MB expression. For example, 247, 30Y and 17D are thought to drive expression in only about one-third of MB neurons or fewer²¹ (M. Mader and M. Heisenberg, personal communication). The neuroanatomical basis of the long sleepers is unclear because substantial expression is observed outside the MB (Supplementary Table S1). Although we cannot exclude a role for the MB in inhibiting sleep, these data indicate that the MB might have a primary function in promoting sleep.

Increases in waking locomotor activity are not uniformly evident in SSL flies. Under temperature cycling, we found that waking activity at 29 °C was not affected in 30Y/UAS shi^{ts1} flies (Fig. 3a). During maintenance at a constant 29 °C, waking activity was also unaffected in some SSL lines (Fig. 3b). In addition to reduced sleep, these lines showed less consolidated sleep. Reduced sleep was not accompanied by significant decreases in sleep bout number (Fig. 3c). However, average bout length (Fig. 3d) and consolidation index, a weighted average bout length (Fig. 3e), were reduced, indicating possibly impaired sleep maintenance. Although 21 °C is nominally permissive, shi^{ts1} effects have been observed at 18 °C (ref. 22), including with 30YGAL4 (data not shown). Nevertheless, consolidation phenotypes were either reduced or absent at 21 °C (Fig. 3f–h).

Previous studies in *Drosophila* have implicated clock genes in regulating sleep^{5,10,23}. We examined two SSL lines during the first four days of temperature cycling (days 3–6) rather than during days 9–12 (Fig. 1a–d). During these days, the circadian contribution to sleep was spread over the day (data not shown). We also examined sleep in SSL lines in a clockless *per⁰¹* background. shi^{ts1} effects were evident under both conditions (Supplementary Fig. S2a–d). In addition, we performed temperature shifts from 21 to 29 °C at different times of day and found that sleep was reduced at 29 °C in GAL4/UAS shi^{ts1} flies at all times (Supplementary Fig. S3a, b). Temperature pulses delivered during a sleep period (ZT 18–24; ZT 0 is time of lights-on and ZT 12 is time of lights-off), in a relatively specific MB line (c309/UAS shi^{ts1}) lacking a 21 °C phenotype, elicits increased sleep loss and subsequent sleep rebound, which is consistent with a sleep regulatory function (Supplementary Fig. S3c).

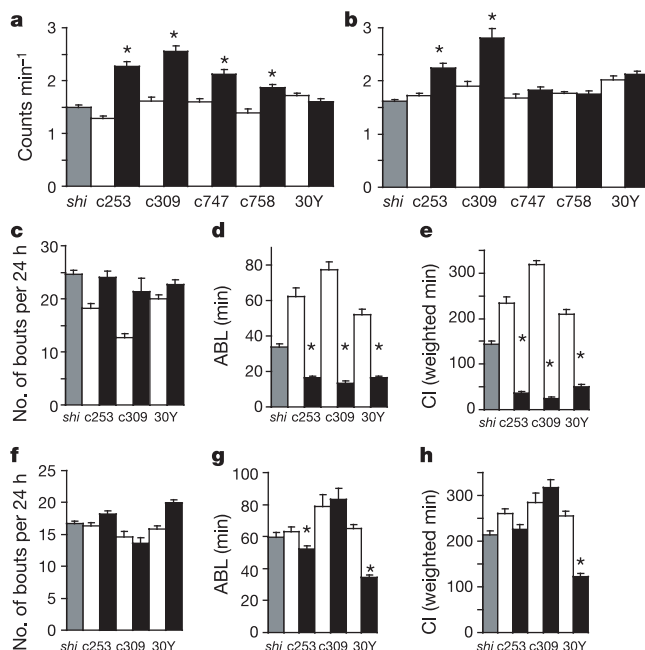


Figure 3 | Sleep intensity is reduced, and activity is not correlated with the short-sleep phenotype. **a, b,** Average waking activity counts per minute during a 29 °C period of temperature cycling (days 9–12) (**a**) and during constant 29 °C conditions (**b**). **c–h,** Sleep intensity: average number of sleep bouts, average sleep bout length (ABL), and sleep consolidation index (CI) under constant 29 °C (**c–e**) or 21 °C (**f–h**) conditions. Grey bars, +/UAS shi^{ts} ; white bars, GAL4/+; black bars, GAL4/UAS shi^{ts} . Asterisk, GAL4/UAS shi^{ts} combination significantly different from both GAL4/+ and +/UAS shi^{ts} controls (one-way ANOVA, $P < 0.05$). In **a**, $n = 33$ –124 and the number of experiments was 3–12; in **b–e**, $n = 24$ –170 and the number of experiments was two to nine. Error bars indicate s.e.m.

In SSL GAL4 lines, expression is observed outside the MB (Fig. 2), and some MB-expressing lines do not have sleep phenotypes (Supplementary Table S1). To assess MB function independently, we used chemical ablation with hydroxyurea. Hydroxyurea fed to larvae 1 h after hatching ablates four neuroblasts that give rise to most MB neurons, and a lateral neuroblast that gives rise to some antennal lobe interneurons²⁴. Hydroxyurea-treated flies showed significant decreases in sleep (Fig. 4a). Although modest increases in sleep bout number were also observed (Fig. 4b), sleep maintenance was primarily affected (Fig. 4c, d), as in GAL4/UAS shi^{ts1} flies (Fig. 3a–c). Although reduced sleep was accompanied by increased waking activity, consistent with previous reports^{20,25} (Fig. 4e), activity and sleep phenotypes are not always correlated. For example, hydroxyurea-treated flies under constant darkness exhibit comparable waking activity to untreated flies in light/dark conditions, whereas their sleep levels are significantly reduced (Fig. 4e). These data provide an independent manipulation of MB function in otherwise genetically identical flies to demonstrate a role for the MB in promoting sleep. We have also observed reduced sleep driving an activated form of protein kinase A (UAS smc^+) by using the MB line c309, a manipulation that does not require temperature changes (data not shown)²⁶. Because this experiment was performed at 25 °C, it argues against a role of heat stress as mediating MB effects on sleep.

The 30Y sleep phenotype is probably not due to its effects outside the MB. 30YGAL4 drives expression in the pars intercerebralis, a locus important for sexually dimorphic aspects of locomotor activity²⁷. Pars intercerebralis GAL4 lines (mai301 and kurs58)²⁸ failed to produce sleep phenotypes comparable to those of 30YGAL4 (data not shown). Additionally, we treated 30YGAL4/UAS shi^{ts1} flies with hydroxyurea but did not observe shi^{ts1} effects

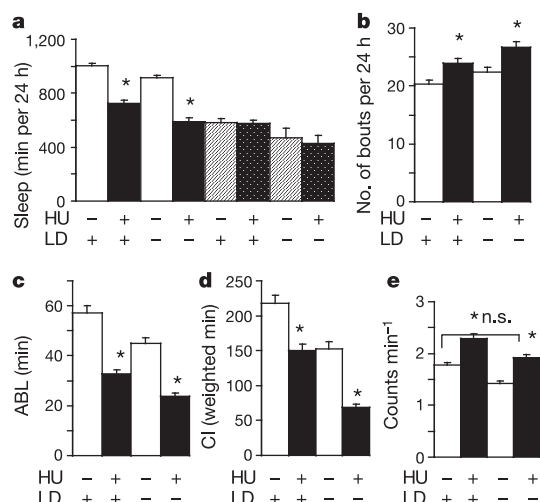


Figure 4 | Mushroom body ablation reduces sleep. **a**, Average total sleep in control and hydroxyurea (HU)-treated wild-type and 30Y/UASShi^{ts1} flies. White bars, wild-type controls; black bars, HU-treated wild type; hatched bars, control 30Y/UASShi^{ts1}; stippled bars, HU-treated 30Y/UASShi^{ts1}. **b–d**, Consolidation: average number of sleep bouts (**b**), average sleep bout length (**c**) and consolidation index (**d**). **e**, Waking activity. Only wild-type flies were used in **b–e**; open bars, controls; filled bars, HU-treated. All flies were assessed at a constant 29 °C in LD and DD. Asterisk, significant difference between control and HU-treated flies (one-way ANOVA, $P < 0.05$). n.s., not significant. For LD, $n = 54$ –84 and the number of experiments was three or four; for DD, $n = 15$ –62 and the number of experiments was two or three. Error bars indicate s.e.m.

in hydroxyurea-treated flies (Fig. 4a). We propose that MB inhibition is largely responsible for the 30Y/UASShi^{ts1} phenotype; MB ablation therefore has no further effect on sleep.

If inhibiting the MB disrupts restorative sleep, then sleep loss may have an adverse consequence on lifespan^{10,29}. Testing SSL and other non-SSL MB lines, we observed significant differences in lifespan curves (data not shown) and, in most cases, mean lifespan for SSL flies (Supplementary Fig. S4a). Although not all SSL lines have a lower mean lifespan, the lifespan for several SSL lines (less than 600 min of sleep per 24 h at 29 °C, is correlated ($r^2 = 0.82$, $P = 0.03$) with the amount of sleep (Supplementary Fig. S4a). At 21 °C, lifespan effects are reduced or absent, with only the SSL line 30Y showing reduced lifespan and sleep (Supplementary Fig. S4b). Moreover, we note reductions in lifespan in hydroxyurea-treated flies relative to their untreated controls (Supplementary Fig. S4c, d). These data indicate that MB-induced sleep reductions may contribute to reduced lifespan.

To determine whether sleep homeostasis was altered in 30Y/UASShi^{ts1} flies, we deprived them of sleep mechanically for 24 h at 29 °C (refs 9, 30) and assayed rebound sleep. When maintained at 29 °C, all flies exhibited a steady increase in sleep (Supplementary Fig. S5a). After detrending (see Supplementary Methods), we found that 30Y/UASShi^{ts1} flies lost less sleep than controls, which was consistent with their reduced sleep at 29 °C (Fig. 1c). Nevertheless, these flies exhibited comparable or increased rebound (Supplementary Fig. S5b), indicating altered sleep homeostasis.

We have uncovered a central role for the MB in sleep regulation. We show that transient MB inhibition by using temperature cycles rapidly inhibits sleep at the restrictive temperature, indicating an active adult function. Persistent inhibition or ablation also reduces sleep, indicating that the MB may promote sleep. The co-localization of sleep and learning centres in *Drosophila* may reflect shared underlying mechanisms, perhaps a role for synaptic plasticity. In this regard, it is of interest that mutants with altered cyclic AMP signalling have disruptions in both learning and sleep^{11,19}. The

discovery of a role for the MB should serve to focus genetic studies to explain the underlying mechanisms of sleep in this model organism.

METHODS

Animals. Flies were raised under a 12 h:12 h light:dark schedule at 25 °C and about 50% humidity. Stocks were provided as follows: GAL4 collection was provided by D. Armstrong through G. Suh; other lines were provided as follows: 30Y/UASShi^{ts1} (A. Presente), pars intercerebralis GAL4 lines (G. Korge), UASShi^{ts1} (T. Kitamoto) and UASMc⁺ (A. Sehgal). Other lines were from the Bloomington Stock Centre. Ablation with hydroxyurea was performed as described previously²⁴. MB ablation was verified with MB expression of GFP.

Sleep assays; measures of sleep and activity. Flies one to four days old were placed into individual 65-mm glass tubes in the *Drosophila* activity monitoring (DAM) system (Trikinetics) under CO₂ anaesthesia. A sleep episode was defined as a 5-min bin of uninterrupted quiescence with the DAM system. Activity counts were summed across all wake bins.

For the temperature cycling assay, after the end of the 12-h light period of the first day of the experiment, flies were kept for 24 h at 21 °C in constant dark (DD). Temperature was then cycled for 14 days (12 h at 29 °C and 12 h at 21 °C in DD). Total sleep and average activity per wake minute were averaged during the 29 or 21 °C period of temperature cycling across four days. Days 9–12 were typically used except where indicated to allow complete circadian re-entrainment.

For the constant-temperature (CT) assay, flies were maintained for five days under 12 h light:12 h dark (LD) conditions followed by seven days of 24 h DD, at either 21 or 29 °C. Total sleep and average activity per wake minute were calculated and averaged across the first three days of LD or DD.

Measures of sleep intensity. All were averaged across three days of LD or DD under constant conditions (21 or 29 °C). Average sleep bout length was calculated by summing all sleep bouts of all lengths (in minutes) and dividing by the total number of sleep bouts. Consolidation index was calculated as a weighted average of sleep bout length, where each sleep bout was weighted according to its duration in minutes. Consolidation index was calculated by summing the square of all sleep bout lengths (in minutes) and dividing by the total amount of sleep. This method reduces the influence of transient awakenings during the sleep phase.

Confocal imaging. Brains of six-day-old male flies were dissected and fixed in 3.7% paraformaldehyde for 40 min. Brains were washed in PBS and in PBS containing 0.3% Triton X-100, incubated and mounted in 80% glycerol and imaged.

Statistical analyses. One-way analysis of variance tests were used to compare averages between different genotypes. $P < 0.05$ was accepted as significant for all analyses.

Received 23 January; accepted 21 March 2006.

- Hendricks, J. C. & Sehgal, A. Why a fly? Using *Drosophila* to understand the genetics of circadian rhythms and sleep. *Sleep* **27**, 334–342 (2004).
- Kitamoto, T. Conditional modification of behaviour in *Drosophila* by targeted expression of a temperature-sensitive *shibire* allele in defined neurons. *J. Neurobiol.* **47**, 81–92 (2001).
- Davis, R. L. Olfactory learning. *Neuron* **44**, 31–48 (2004).
- Shaw, P. J., Cirelli, C., Greenspan, R. J. & Tononi, G. Correlates of sleep and waking in *Drosophila melanogaster*. *Science* **287**, 1834–1837 (2000).
- Hendricks, J. C. *et al.* Rest in *Drosophila* is a sleep-like state. *Neuron* **25**, 129–138 (2000).
- Hendricks, J. C., Kirk, D., Panckeri, K., Miller, M. S. & Pack, A. I. Modafinil maintains waking in the fruit fly *Drosophila melanogaster*. *Sleep* **26**, 139–146 (2003).
- van Swinderen, B., Nitz, D. A. & Greenspan, R. J. Uncoupling of brain activity from movement defines arousal states in *Drosophila*. *Curr. Biol.* **14**, 81–87 (2004).
- Nitz, D. A., van Swinderen, B., Tononi, G. & Greenspan, R. J. Electrophysiological correlates of rest and activity in *Drosophila melanogaster*. *Curr. Biol.* **12**, 1934–1940 (2002).
- Huber, R. *et al.* Sleep homeostasis in *Drosophila melanogaster*. *Sleep* **27**, 628–639 (2004).
- Shaw, P. J., Tononi, G., Greenspan, R. J. & Robinson, D. F. Stress response genes protect against lethal effects of sleep deprivation in *Drosophila*. *Nature* **417**, 287–291 (2002).
- Hendricks, J. C. *et al.* A non-circadian role for cAMP signaling and CREB activity in *Drosophila* rest homeostasis. *Nature Neurosci.* **4**, 1108–1115 (2001).
- Cirelli, C. *et al.* Reduced sleep in *Drosophila* Shaker mutants. *Nature* **434**, 1087–1092 (2005).
- Kosaka, T. & Ikeda, K. Possible temperature-dependent blockage of synaptic vesicle recycling induced by a single gene mutation in *Drosophila*. *J. Neurobiol.* **14**, 207–225 (1983).

14. van der Bliek, A. M. & Meyerowitz, E. M. Dynamin-like protein encoded by the *Drosophila shibire* gene associated with vesicular traffic. *Nature* **351**, 411–414 (1991).
15. Chen, M. S. *et al.* Multiple forms of dynamin are encoded by *shibire*, a *Drosophila* gene involved in endocytosis. *Nature* **351**, 583–586 (1991).
16. Koenig, J. H., Saito, K. & Ikeda, K. Reversible control of synaptic transmission in a single gene mutant of *Drosophila melanogaster*. *J. Cell Biol.* **96**, 1517–1522 (1983).
17. Armstrong, J. D. & Kaiser, K. Flytrap lines. (www.fly-trap.org) (1996).
18. Pascual, A. & Preat, T. Localization of long-term memory within the *Drosophila* mushroom body. *Science* **294**, 1115–1117 (2001).
19. Zars, T., Fischer, M., Schulz, R. & Heisenberg, M. Localization of a short-term memory in *Drosophila*. *Science* **288**, 672–675 (2000).
20. Martin, J. R., Ernst, R. & Heisenberg, M. Mushroom bodies suppress locomotor activity in *Drosophila melanogaster*. *Learn. Mem.* **5**, 179–191 (1998).
21. Schwaerzel, M., Heisenberg, M. & Zars, T. Extinction antagonizes olfactory memory at the subcellular level. *Neuron* **35**, 951–960 (2002).
22. Sapp, R. J., Christianson, J. & Stark, W. S. Turnover of membrane and opsin in visual receptors of normal and mutant *Drosophila*. *J. Neurocytol.* **20**, 597–608 (1991).
23. Hendricks, J. C. *et al.* Gender dimorphism in the role of cycle (BMAL1) in rest, rest regulation, and longevity in *Drosophila melanogaster*. *J. Biol. Rhythms* **18**, 12–25 (2003).
24. de Belle, J. S. & Heisenberg, M. Associative odor learning in *Drosophila* abolished by chemical ablation of mushroom bodies. *Science* **263**, 692–695 (1994).
25. Helfrich-Forster, C., Wulf, J. & de Belle, J. S. Mushroom body influence on locomotor activity and circadian rhythms in *Drosophila melanogaster*. *J. Neurogenet.* **16**, 73–109 (2002).
26. Li, W., Ohlmeyer, J. T., Lane, M. E. & Kalderon, D. Function of protein kinase A in hedgehog signal transduction and *Drosophila* imaginal disc development. *Cell* **80**, 553–562 (1995).
27. Belgacem, Y. H. & Martin, J. R. Neuroendocrine control of a sexually dimorphic behaviour by a few neurons of the pars intercerebralis in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **99**, 15154–15158 (2002).
28. Siegmund, T. & Korge, G. Innervation of the ring gland of *Drosophila melanogaster*. *J. Comp. Neurol.* **431**, 481–491 (2001).
29. Rechtschaffen, A., Gilliland, M. A., Bergmann, B. M. & Winter, J. B. Physiological correlates of prolonged sleep deprivation in rats. *Science* **221**, 182–184 (1983).
30. Cirelli, C. Searching for sleep mutants of *Drosophila melanogaster*. *Bioessays* **25**, 940–949 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Villar, A. Schroeder, B. Finn, D. Hanrahan, A. Majeed and A. Phillips for assistance with PCR genotyping (M.V.), confocal imaging (A.S., B.F.), initial screening (D.H., A.M.) and designing data analysis software (A.P.); C. Cirelli and G. Tononi and their laboratories for advice on mechanical sleep deprivation; P. Shaw, T. Zars, M. Rosbash and E. Smith for comments; and B. Joiner and A. Sehgal for communicating results before publication. This work was supported by a Burroughs Wellcome Career Award in the Biomedical Sciences and by the NIH (R.A.).

Author Contributions J.L.P. completed all experiments and analyses, with assistance on lifespan, PCR genotyping and general fly maintenance from J.J.McG., and on the development and application of behaviour data analysis software for measures of sleep intensity (the *Drosophila* Activity Monitor Data Crunching Macro) from K.P.K. J.L.P. and R.A. wrote the manuscript.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.A. (r-allada@northwestern.edu).

Sleep in *Drosophila* is regulated by adult mushroom bodies

William J. Joiner¹, Amanda Crocker¹, Benjamin H. White³ & Amita Sehgal^{1,2}

Sleep is one of the few major whole-organ phenomena for which no function and no underlying mechanism have been conclusively demonstrated. Sleep could result from global changes in the brain during wakefulness or it could be regulated by specific loci that recruit the rest of the brain into the electrical and metabolic states characteristic of sleep. Here we address this issue by exploiting the genetic tractability of the fruitfly, *Drosophila melanogaster*, which exhibits the hallmarks of vertebrate sleep^{1–4}. We show that large changes in sleep are achieved by spatial and temporal enhancement of cyclic-AMP-dependent protein kinase (PKA) activity specifically in the adult mushroom bodies of *Drosophila*. Other manipulations of the mushroom bodies, such as electrical silencing, increasing excitation or ablation, also alter sleep. These results link sleep regulation to an anatomical locus known to be involved in learning and memory.

Sleep duration is inversely related to PKA activity in flies⁵. To determine whether specific brain loci regulate sleep, we used the GAL4/UAS (upstream activating sequence) system⁶ to express a catalytic subunit of PKA in various regions of the fly brain. We first expressed PKA under the control of the RU486-inducible pan-neuronal driver *elavGeneSwitch*⁷. Restricting the expression of PKA to adult neurons decreased daily sleep, supporting earlier studies with mutants such as *dunce*⁵ that increase PKA levels, and showing that PKA directly regulates sleep rather than a developmental process that might affect sleep (Fig. 1a). We next expressed PKA under the control of different GAL4 drivers and examined the changes in total daily sleep in the different driver/transgene combinations relative to driver/background and background/transgene controls (Fig. 1b). When both controls were taken into account, the expression of PKA by only two drivers led to changes in sleep that exceeded 2 s.d. These were 201Y, which increased sleep by $75 \pm 3\%$ and $93 \pm 4\%$ respectively, and c309, which decreased sleep by $46 \pm 11\%$ and $43 \pm 14\%$ per day compared with the two sets of controls. Changes in sleep caused by all other GAL4 drivers remained within 1 s.d. of the mean.

We next examined whether activity levels during wake periods were affected by the 201Y and c309 drivers. Many GAL4 driver/UAS-PKA lines were hypoactive, but line 201Y had normal waking activity (Fig. 1c). Similarly, activity normalized to waking time in c309 was not significantly higher in PKA-driven animals than in either control (Fig. 1c), indicating that c309 was not hyperactive. We conclude that the sleep phenotypes of animals expressing PKA under control of the 201Y and c309 drivers are not associated with abnormal waking activity. Interestingly, both these drivers are known to be expressed in the mushroom bodies (MBs)⁸, a brain region implicated in associative learning.

Given the strong, yet opposite, effects that 201Y and c309 had on sleep, we further characterized their expression patterns in the fly brain by crossing them into animals bearing a UAS transgene for

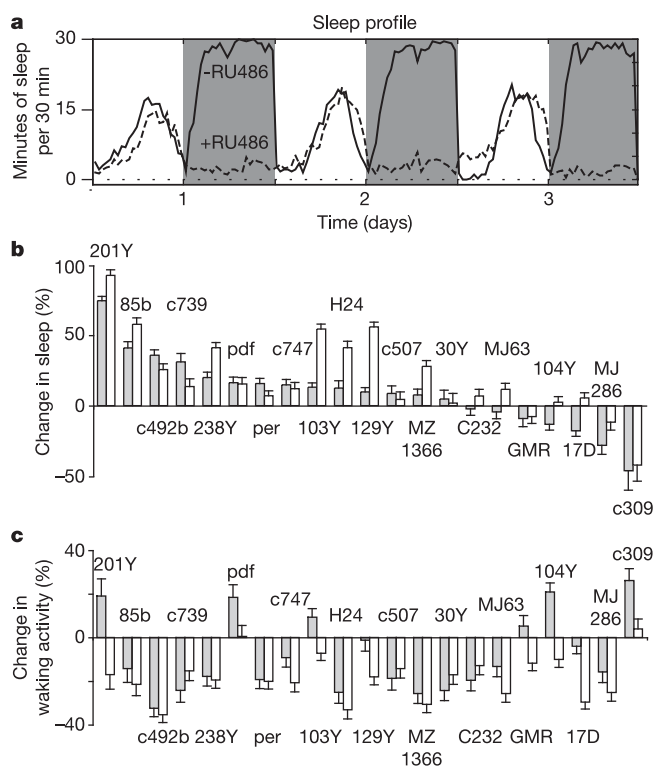


Figure 1 | Expression of PKA in various regions of the fly brain affects sleep differentially. **a**, The pan-neuronal driver *elavGeneSwitch*⁷ was used to express UAS-mc*, which contains a constitutively active subunit of PKA downstream of an upstream activating sequence (UAS)³⁰. Sleep, defined as 5 min of immobility, was assayed in flies in which expression of PKA was induced with 500 μ M RU486 (dashed line) and in uninduced controls (solid line). **b**, Progeny were collected from crosses between 21 GAL4 drivers and *y w;UAS-mc**. For controls, each GAL4 line was crossed to *y w* flies (the background for the UAS-transgenic line), and *y w;UAS-mc** was crossed to *w¹¹¹⁸/y w* (the background for the GAL4 drivers). Daily sleep was averaged over 4 days. Sleep from control animals was subtracted from the experimental group, then divided by the amount of control sleep to calculate net percentage changes. Grey bars, sleep in *w¹¹¹⁸/y w* flies carrying UAS-mc* and a GAL4 transgene relative to sleep in *w¹¹¹⁸/y w* flies carrying the respective GAL4 transgene alone; white bars, sleep in *w¹¹¹⁸/y w* flies carrying UAS-mc* and a GAL4 transgene relative to sleep in *w¹¹¹⁸/y w* flies carrying only UAS-mc*. **c**, For each of the crosses in **b**, average daily waking activity (beam crossings) was divided by average total wake time. Differences in waking activity between GAL4/UAS-mc* and control flies were calculated as in **b**. For each group, $n \geq 25$. Where errors are shown, they are s.e.m.

¹Center for Sleep and Respiratory Neurobiology and ²Howard Hughes Medical Institute, University of Pennsylvania Medical School, Philadelphia, Pennsylvania 19104, USA.

³Laboratory of Molecular Biology, National Institute of Mental Health, Bethesda, Maryland 20892, USA.

green fluorescent protein (GFP). We found that 201Y is expressed largely in the γ lobes and the core region of the α/β lobes of the MBs, whereas c309 is expressed in the α/β and γ lobes but not in the core region of the α/β lobes (Fig. 2a, b, and Supplementary Fig. S1). This differential expression pattern within the MBs indicates that PKA might affect the regulation of sleep by the MBs in both a positive and a negative fashion by using anatomically distinct classes of neurons. Consistent with this notion of heterogeneous cell types within the MBs^{9–11}, some MB drivers, such as 30Y and 238Y, promoted sleep during the day but inhibited sleep during the night, leading to only marginal overall changes in daily sleep (Fig. 1, and Supplementary Fig. S2). This effect was not observed with any driver that was expressed exclusively outside the MBs (Fig. 2c and Supplementary Table 1). A small increase in daytime sleep was also frequently produced by the pan-neuronal *elav*GeneSwitch driver, which decreased overall sleep (data not shown, and Fig. 1a). The expression patterns of 238Y and 30Y overlap those of 201Y and c309, supporting the idea that 238Y and 30Y are expressed in both sleep-promoting and sleep-inhibiting areas (see Supplementary Table S1 and Supplementary Fig. S1 for expression patterns of drivers used in this study).

To test the hypothesis that PKA expression in MBs regulates adult sleep, we expressed the PKA transgene under the control of an RU486-activatable MB GAL4 driver, P{MB-Switch}¹². We confirmed selective expression of this driver in the MBs by coupling it to a GFP reporter, and found inducible expression in the MBs (Fig. 2d, e, and Supplementary Fig. S1). Sleep was significantly reduced in response to RU486 in MB-Switch/PKA animals (Fig. 3a, b) but was unaffected by the drug in control animals harbouring either the driver or the transgene alone (Fig. 3c). Thus, PKA overexpressed preferentially in specific neurons of adult MBs is sufficient to reduce sleep.

Next we compared sleep structure in the hyposomnolent animals with that of controls. In both MB-Switch/PKA animals and c309/PKA animals, loss of sleep was caused by a shortened sleep bout duration without a concomitant increase in the sleep bout number (Fig. 3d, e, and Supplementary Fig. S3a, b). The underlying cause of reduced sleep in both sets of animals therefore seems to be impaired sleep need, because the alternative—normal sleep need, but an inability to

maintain the sleep state—would be expected to produce an increase in sleep bout number. In contrast, in 201Y sleep bout duration remained unchanged (Supplementary Fig. S3c, d).

We then asked whether the reduction of sleep in MB-Switch/PKA

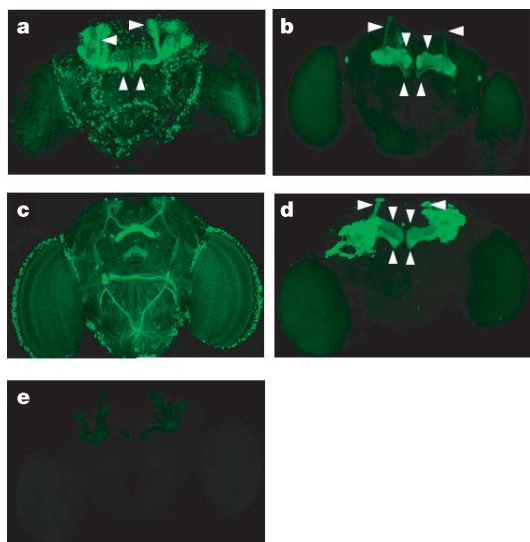


Figure 2 | Localization of GAL4-dependent brain expression. A nuclear-targeted GFP (GFPn) was expressed under the control of different GAL4 drivers. **a**, c309/GFPn; upper arrows refer to α lobes and lower arrows to β lobes of MBs. **b**, 201Y/GFPn; upper, middle and lower arrows point to α , γ and β lobes, respectively, within MBs. **c**, 104Y/GFPn; the fan-shaped body of the central complex and cells rimming the optic lobes and anterior brain are illuminated. **d**, P{MB-Switch}/GFPn with RU486 in the food; labelling is as in **b**. **e**, P{MB-Switch}/GFPn without RU486 in the food.

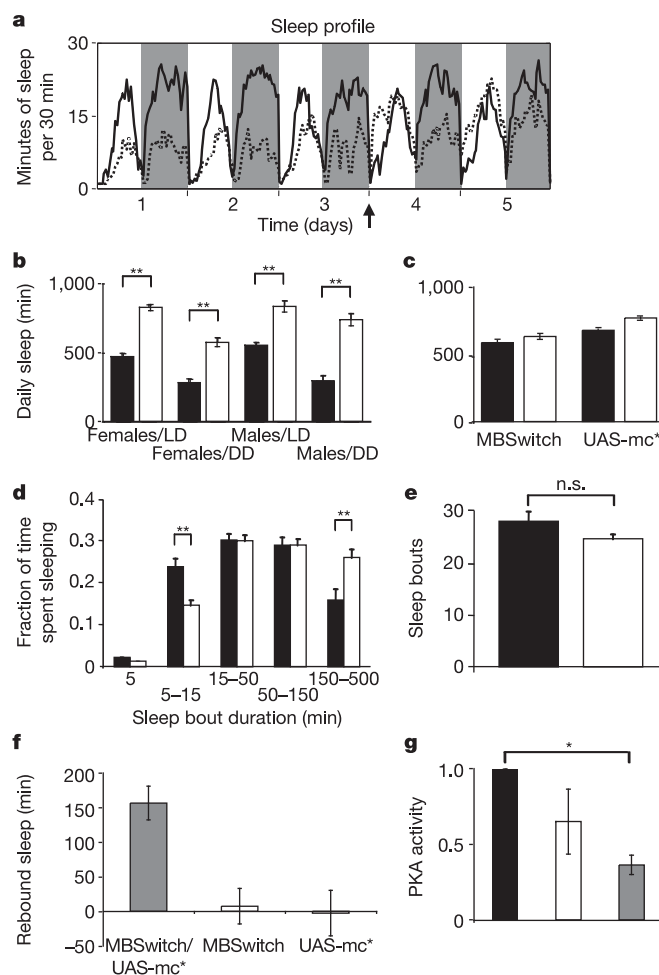


Figure 3 | Inducible expression of PKA in the MBs leads to decreased sleep

bout duration and accumulation of sleep debt. **a**, Sleep was monitored in *w¹¹¹⁸/y w; UAS-mc⁺/+; P{MB-Switch}/+* females in the presence (dashed line) or absence (solid line) of 500 μ M RU486 ($n \geq 16$ for each group). At the end of day 3, animals were transferred to tubes lacking RU486 (arrow). Grey bars represent 12-h dark periods. **b**, Sleep was measured for females in LD ($n = 61$ or 62) and in constant darkness (DD, $n = 28$ –30), and for males in LD ($n = 45$ or 46) and in DD ($n = 15$ or 16). Black bars represent animals fed RU486; white bars are controls. Plotted values represent averages over three days. **c**, *w¹¹¹⁸/y w; P{MB-Switch}/+* or *w¹¹¹⁸/y w; UAS-mc⁺/+* female controls were placed in LD for 3 days with (black bars) or without (white bars) RU486 in their food. **d**, Sleep bouts in LD were binned according to duration for P{MB-Switch}/UAS-mc⁺ females with (black bars) or without (white bars) RU486 ($n = 43$ –45 for each group). **e**, Sleep bout number in LD is similar for P{MB-Switch}/UAS-mc⁺ females treated with RU486 (black bars) to that in genotypically identical uninduced animals (white bars). $n = 43$ –45 for each group. **f**, Flies were monitored for four days in the presence or absence of 500 μ M RU486 and then transferred to tubes with no drug at diurnal time 0 (ZT0; when the light was turned on). Rebound sleep during the following 12 hours was determined by subtracting sleep in the 'no drug to no drug' group from sleep in 'drug to no drug' animals. $n > 64$ for *w¹¹¹⁸/y w; UAS-mc⁺/+; P{MB-Switch}/+* experimental animals; $n > 30$ for *w¹¹¹⁸/y w; P{MB-Switch}/+* and *w¹¹¹⁸/y w; UAS-mc⁺/+* controls. **g**, *elav*-GeneSwitch/UAS-mc⁺ females were switched at ZT0 from drug to drug (black bar), from no drug to no drug (white bar) or from drug to no drug (that is rebound conditions, grey bar). After 2 h, PKA activity was measured in heads from each group. Significance was determined with a one-way analysis of variance. Asterisk, $P < 0.05$; two asterisks, $p < 0.01$; n.s., $p > 0.05$ by unpaired *t*-test. Where errors are shown, they are s.e.m.

animals was due to an impaired accrual of a sleep-inducing signal. If this were so, then a hallmark of sleep, homeostatic rebound—sleep that exceeds baseline to compensate for lost sleep—should not occur on relief of induced PKA expression. However, when RU486 was withdrawn after about three days of sleep deprivation, an average rebound of 156 ± 38 min was observed (Fig. 3a, f). This is a robust rebound, comparable to that produced when genetically identical but uninduced flies were submitted to a standard 12 h of mechanical deprivation (137 ± 26 min; Supplementary Fig. S4). Behavioural rebound was also observed in animals expressing *elavGeneSwitch*-driven PKA, after withdrawal of RU486 (data not shown), and was accompanied by a decrease in PKA activity in fly heads (Fig. 3g). Rebound after withdrawal of RU486 indicates that PKA might not prevent the accrual of sleep-promoting signals but might suppress homeostatic output.

To determine whether PKA affects sleep by regulating synaptic output in MB neurons, we inducibly expressed either of two K^+ channels, Kir2.1 (ref. 13) or EKO¹⁴, under the control of the MB-Switch driver. Such transgenic expression should suppress action-potential firing by hyperpolarizing neurons and decreasing membrane resistance, thus leading to reduced synaptic transmission. We found that induction of either Kir2.1 or EKO caused a significant increase in sleep (Fig. 4a, b, and Supplementary Fig. S5). Because the opposite was observed with PKA expression in the same neurons,

it indicates that PKA might decrease sleep by increasing either excitability or synaptic transmission. To address this issue further, we inducibly expressed a sodium channel (NaChBac), which depolarizes neurons and increases excitability¹⁵. When expressed under the control of the MB-Switch driver, the sodium channel caused a decrease in sleep (Fig. 4c, d), similar to that produced by PKA, confirming that PKA increases the output of these neurons.

The MB-Switch driver is expressed in a subpopulation of MB neurons similar to those labelled by c309 (Supplementary Fig. S1), and both drivers had sleep-inhibiting effects. As noted above, this pattern of expression differed from that of other drivers (Supplementary Fig. S1), which had sleep-promoting or bidirectional effects on sleep (Fig. 1b, and Supplementary Fig. S2), thus leading us to propose that the MBs contain sleep-inhibiting and sleep-promoting neurons. To determine the overall effect of MBs on sleep, we ablated them with hydroxyurea and examined sleep and activity in adult flies. Consistent with previous reports^{16,17} was our observation of an overall increase in activity (Supplementary Fig. S6a). However, normalization of this activity to waking time indicates that the phenotype derives less from hyperactivity than from a reduction in sleep (Supplementary Fig. S6b, c). Even so, the reduction in sleep was much less than that seen with other manipulations of the MBs (see above) or in short-sleep mutants such as *minisleep*¹⁸. This supports our conclusion that MBs exert both positive and negative influences on sleep that are integrated to produce the overt behavioural state. Our model (Supplementary Fig. S7) takes into account our results; notably the integrator downstream of the MBs promotes activity in the default state. Thus, when MBs are ablated the overall effect is increased wakefulness.

Opposing effects of the c309 and 201Y drivers are also observed in a different behavioural model. They parallel MB-dependent changes in brain activity during the sleep/wake cycle that are associated with salience, a behavioural trait that may correspond to arousal. Consistent with our data was the observation that reducing synaptic transmission using the c309 driver inhibited salience, whereas the 201Y driver in the same type of experiment yielded no change. We would predict increased arousal with 201Y, but in those experiments the animals were already awake⁴.

Because MBs receive and transduce considerable sensory, particularly olfactory, input to the fly brain, we speculate that they promote arousal or sleep by allowing or inhibiting the throughput of sensory information. In addition, given the major function that MBs have in regulating plasticity in the fly brain, it is likely that this is linked to their role in sleep. In mammals, sleep deprivation suppresses the performance of learned tasks, and sleep permits memory consolidation¹⁹. Sleep and sleep deprivation also differentially affect cortical synaptic plasticity²⁰. In *Drosophila*, MBs participate in the consolidation or retrieval of memories involving olfactory cues^{21–24}, courtship conditioning^{8,25} and context-dependent visual cues^{26,27} by mechanisms that include cAMP signalling. Distinct anatomical regions of the MBs have been shown to be important for at least some forms of memory^{23,28}, as we have now also shown for sleep. Thus, memory and sleep may involve similar molecular pathways (cAMP signalling) and anatomical regulatory loci (MBs).

METHODS

Sleep assays. Female flies 3–7 days old were placed in 65 mm $\times$ 5 mm glass tubes containing 5% sucrose/2% agarose with or without 500 μ M RU486. Flies were acclimated for about 36 h at 25 °C in 12 h light/12 h dark (LD) conditions. Locomotor activity was collected in LD with DAMS monitors (Trikinetics) as described previously¹⁵. Sleep was measured as bouts of 5 min of inactivity, as described previously²⁹, using a moving window over 30-s intervals. In some experiments RU486 was removed at the time points indicated in figure legends. **Localization of GAL4 expression.** Many of the GAL4 lines from Fig. 1 were crossed to a fly line with a transgene encoding GFP fused to a nuclear localization signal (GFPn). Brains were dissected and fixed for 20–30 min in 4% paraformaldehyde in PBS before photography of serial z-sections of whole mounts with the use of confocal microscopy. In some cases, images were deconvolved to allow

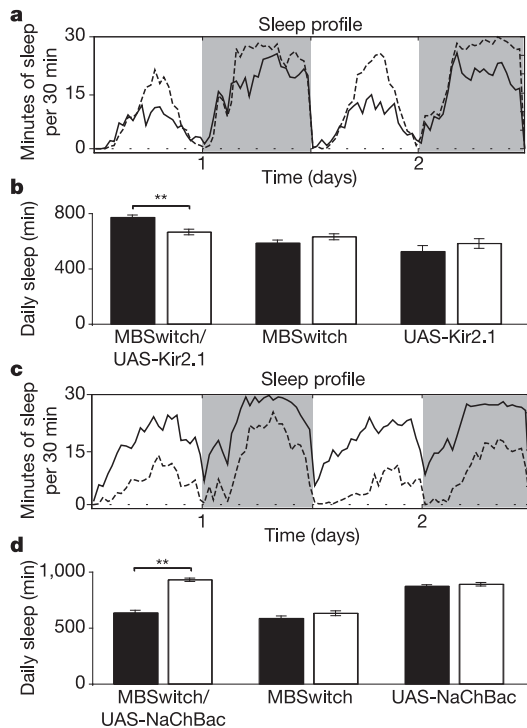


Figure 4 | Decreasing and increasing excitability in a subset of MB neurons have opposite effects on sleep. **a**, In P{MB-Switch}/UAS-Kir2.1 animals in which K^+ channel expression in the MBs was induced by RU486 (dashed line), sleep exceeded that of control animals of identical genotype (solid line). **b**, RU486 significantly increased sleep in MBSwitch/Kir2.1 females (black bars); white bars are uninduced controls. From left to right: for w^{1118}/y w; MBSwitch/Kir2.1, $n = 34$ –47; for w^{1118}/y w; MBSwitch/+, $n = 30$ or 31; for w^{1118}/y w; UAS-Kir2.1/+, $n = 11$ –26. **c**, In P{MB-Switch}/UAS-NaChBac flies in which Na^+ channel expression in the MBs was induced by RU486 (dashed line), sleep was significantly suppressed in comparison with control animals of identical genotype (solid line). **d**, RU486 significantly suppressed sleep in MBSwitch/UAS-NaChBac females. From left to right: for w^{1118}/y w; UAS-NaChBac/+; MBSwitch/+, $n = 46$ or 47; for w^{1118}/y w; MBSwitch/+, $n = 31$; for w^{1118}/y w; UAS-NaChBac/+, $n = 59$ or 60. Two asterisks, $P < 0.01$ by unpaired t -test. Where errors are shown, they are s.e.m.

virtual transverse sections to be taken through MB peduncles.

PKA assays. After 1 week in LD, elav-GeneSwitch/UAS-mc* females were switched at ZT0 from drug to drug, from no drug to no drug or from drug to no drug. After 2 h, heads were isolated and homogenized on ice in buffer consisting of (in mM): 10 HEPES pH 7.5, 100 KCl, 1 EDTA, 5 dithiothreitol, 5 phenylmethylsulphonyl fluoride, 10% glycerol, 0.1% Triton X-100 and one tablet of Complete (Promega). Debris was pelleted for 5 min at 4 °C in a microcentrifuge and discarded. A protein assay (Dc kit; Bio-Rad) was performed on supernatant from each sample, and homogenization buffer was added as necessary to make each sample equal in concentration. PKA activity was measured as ATP-dependent quenching of fluorescent kemptide in accordance with the manufacturer's instructions (IQ assay; Pierce) and normalized to values determined for animals maintained on RU486 throughout the experiment.

Hydroxyurea-dependent ablation of MBs. Mated Canton-S flies were placed on standard grape juice/agar plates to induce egg-laying over a 1-h period. On the next day, first-instar larvae were collected within 1 h of hatching and placed in 50% yeast paste in the presence or absence of 50 mg ml⁻¹ hydroxyurea for 4 h before being transferred to standard food vials, as described previously²¹. After eclosion, adult male flies were aged for one week and then placed in glass tubes to measure sleep/activity patterns over a three-day period in LD.

Received 25 January; accepted 12 April 2006.

- Hendricks, J. C. *et al.* Rest in *Drosophila* is a sleep-like state. *Neuron* **25**, 129–138 (2000).
- Shaw, P. J., Cirelli, C., Greenspan, R. J. & Tononi, G. Correlates of sleep and waking in *Drosophila melanogaster*. *Science* **287**, 1834–1837 (2000).
- Nitz, D. A., van Swinderen, B., Tononi, G. & Greenspan, R. J. Electrophysiological correlates of rest and activity in *Drosophila melanogaster*. *Curr. Biol.* **12**, 1934–1940 (2002).
- van Swinderen, B. & Greenspan, R. J. Salience modulates 20–30 Hz brain activity in *Drosophila*. *Nature Neurosci.* **6**, 579–586 (2003).
- Hendricks, J. C. *et al.* A non-circadian role for cAMP signaling and CREB activity in *Drosophila* rest homeostasis. *Nature Neurosci.* **4**, 1108–1115 (2001).
- Brand, A. H. & Perrimon, N. Targeted expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Osterwalder, T., Yoon, K. S., White, B. H. & Keshishian, H. A conditional tissue-specific transgene expression system using inducible GAL4. *Proc. Natl Acad. Sci. USA* **98**, 12596–12601 (2001).
- Joiner, M. A. & Griffith, L. C. Mapping of the anatomical circuit of CaM kinase-dependent courtship conditioning in *Drosophila*. *Learn. Mem.* **6**, 177–192 (1999).
- Armstrong, J. D., de Belle, J. S., Wang, Z. & Kaiser, K. Metamorphosis of the mushroom bodies; large-scale rearrangements of the neural substrates for associative learning and memory in *Drosophila*. *Learn. Mem.* **5**, 102–114 (1998).
- Zhu, S., Chiang, A. S. & Lee, T. Development of the *Drosophila* mushroom bodies: elaboration, remodeling and spatial organization of dendrites in the calyx. *Development* **130**, 2603–2610 (2003).
- Crittenden, J. R., Skoulakis, E. M., Han, K. A., Kalderon, D. & Davis, R. L. Tripartite mushroom body architecture revealed by antigenic markers. *Learn. Mem.* **5**, 38–51 (1998).
- Mao, Z., Roman, G., Zong, L. & Davis, R. L. Pharmacogenetic rescue in time and space of the rutabaga memory impairment using Gene-Switch. *Proc. Natl Acad. Sci. USA* **101**, 198–203 (2004).
- Baines, R. A., Uhler, J. P., Thompson, A., Sweeney, S. T. & Bate, M. Altered electrical properties in *Drosophila* neurons developing without synaptic transmission. *J. Neurosci.* **21**, 1523–1531 (2001).
- White, B. H. *et al.* Targeted attenuation of electrical activity in *Drosophila* using a genetically modified K⁺ channel. *Neuron* **31**, 699–711 (2001).
- Luan, H. *et al.* Functional dissection of a neuronal network required for cuticle tanning and wing expansion in *Drosophila*. *J. Neurosci.* **26**, 573–584 (2006).
- Martin, J. R., Ernst, R. & Heisenberg, M. Mushroom bodies suppress locomotor activity in *Drosophila melanogaster*. *Learn. Mem.* **5**, 179–191 (1998).
- Helfrich-Forster, C., Wulf, J. & de Belle, J. S. Mushroom body influence on locomotor activity and circadian rhythms in *Drosophila melanogaster*. *J. Neurogenet.* **16**, 73–109 (2002).
- Cirelli, C. *et al.* Reduced sleep in *Drosophila* Shaker mutants. *Nature* **434**, 1087–1092 (2005).
- Walker, M. P. & Stickgold, R. Sleep-dependent learning and memory consolidation. *Neuron* **44**, 121–133 (2004).
- Frank, M. G., Issa, N. P. & Stryker, M. P. Sleep enhances plasticity in the developing visual cortex. *Neuron* **30**, 275–287 (2001).
- de Belle, J. S. & Heisenberg, M. Associative odor learning in *Drosophila* abolished by chemical ablation of mushroom bodies. *Science* **263**, 692–695 (1994).
- Dubnau, J., Grady, L., Kitamoto, T. & Tully, T. Disruption of neurotransmission in *Drosophila* mushroom body blocks retrieval but not acquisition of memory. *Nature* **411**, 476–480 (2001).
- McGuire, S. E., Le, P. T. & Davis, R. L. The role of *Drosophila* mushroom body signaling in olfactory memory. *Science* **293**, 1330–1333 (2001).
- Zars, T., Fischer, M., Schulz, R. & Heisenberg, M. Localization of a short-term memory in *Drosophila*. *Science* **288**, 672–675 (2000).
- McBride, S. M. *et al.* Mushroom body ablation impairs short-term memory and long-term memory of courtship conditioning in *Drosophila melanogaster*. *Neuron* **24**, 967–977 (1999).
- Liu, L., Wolf, R., Ernst, R. & Heisenberg, M. Context generalization in *Drosophila* visual learning requires the mushroom bodies. *Nature* **400**, 753–756 (1999).
- Tang, S. & Guo, A. Choice behavior of *Drosophila* facing contradictory visual cues. *Science* **294**, 1543–1547 (2001).
- Pascual, A. & Preat, T. Localization of long-term memory within the *Drosophila* mushroom body. *Science* **94**, 1115–1117 (2001).
- Huber, R. *et al.* Sleep homeostasis in *Drosophila melanogaster*. *Sleep* **27**, 628–639 (2004).
- Li, W., Ohlmeyer, J. T., Lane, M. E. & Kalderon, D. Function of protein kinase A in hedgehog signal transduction and *Drosophila* imaginal disc development. *Cell* **80**, 553–562 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Roman and R. Davis for P{MB-Switch}, and L. Griffith, T. Kitamoto, R. Greenspan, J. D. Armstrong, K. Kaiser, R. Allada and C. Helfrich-Forster for other GAL4 lines; J. Hendricks for advice on studying fly sleep; E. Friedman, S. Harbison, K. Ho, K. Koh, S. Sathyanarayanan, M. Wu, Q. Yuan and X. Zheng for critical comments on the manuscript; and K. Liu for assistance with animal maintenance. We also thank J. Pitman and R. Allada for communicating results before publication. W.J.J. and A.C. were supported by training grants to the Center for Sleep and Respiratory Neurobiology at the University of Pennsylvania, and W.J.J. was also supported by a Pickwick Fellowship from the National Sleep Foundation. Partial support was provided by a programme project grant from the NIA. B.H.W. was supported by the Intramural Research Program of the NIH, NIMH.

Author Contributions W.J.J. conducted the behavioural experiments, wrote the sleep analysis software, analysed the sleep data, prepared fly brains and heads for confocal microscopy and PKA assays, and co-wrote the manuscript. A.C. assisted with the behavioural experiments and sleep analysis, prepared fly brains for imaging, performed all confocal microscopy, and provided feedback on the manuscript. B.H.W. provided NaChBac transgenic animals and feedback on the manuscript. A.S. conceived the project, provided guidance and critical interpretation during the project, co-wrote the manuscript and provided financial support.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.S. (amita@mail.med.upenn.edu).

Modulation of intracortical synaptic potentials by presynaptic somatic membrane potential

Yousheng Shu¹, Andrea Hasenstaub¹, Alvaro Duque¹, Yuguo Yu¹ & David A. McCormick¹

Traditionally, neuronal operations in the cerebral cortex have been viewed as occurring through the interaction of synaptic potentials in the dendrite and soma, followed by the initiation of an action potential, typically in the axon^{1,2}. Propagation of this action potential to the synaptic terminals is widely believed to be the only form of rapid communication of information between the soma and axonal synapses, and hence to postsynaptic neurons. Here we show that the voltage fluctuations associated with dendrosomatic synaptic activity propagate significant distances along the axon, and that modest changes in the somatic membrane potential of the presynaptic neuron modulate the amplitude and duration of axonal action potentials and, through a Ca^{2+} -dependent mechanism, the average amplitude of the postsynaptic potential evoked by these spikes. These results indicate that synaptic activity in the dendrite and soma controls not only the pattern of action potentials generated, but also the amplitude of the synaptic potentials that these action potentials initiate in local cortical circuits, resulting in synaptic transmission that is a mixture of triggered and graded (analogue) signals.

To examine the properties of monosynaptic excitatory connections within the cortex, we obtained whole-cell recordings from nearby (within 60 μm) and synaptically connected pairs of layer 5 pyramidal cells in ferret prefrontal cortical slices (Fig. 1, $n = 28$ paired recordings; Supplementary Fig. 1). As expected, activation of an action potential in the presynaptic neuron with a short (1 ms) depolarizing current pulse resulted in an excitatory postsynaptic potential (EPSP; Fig. 1b). However, we found that steady depolarization of the soma of the presynaptic neuron by approximately 10–15 mV from the resting membrane potential (average -62 mV) to near the firing threshold (average -48 mV) resulted in a substantial and statistically significant increase in the average EPSP amplitude evoked in the postsynaptic cell (18/28 synaptic connections in 28 unique paired recordings; Fig. 1b, e), even though this depolarization had no significant effect on the resting membrane potential of the postsynaptic cell. The median significant enhancement was 29% (ranging from 12 to 69%, with one outlier at 194%; Fig. 1e) and was reversible (Fig. 1b).

We observed the enhancing effect of presynaptic somatic membrane potential even if the hyperpolarized and depolarized periods were 10–20 s in duration each and intermixed (Fig. 1c, d; $n = 19/37$ synaptic connections in 33 pairs tested). We examined the kinetics of the enhancing effect of presynaptic somatic depolarization on EPSP amplitude in paired recordings and found that the enhancement seemed to have both rapid (<1 s) and slower (>1 s) components (Fig. 1f). The disfacilitation upon subsequent hyperpolarization exhibited a slow decrease that could be well fitted ($r^2 = 0.92$) by an exponential function with $\tau = 4.3$ s ($n = 10$, Fig. 1f; see also Supplementary Fig. 2).

We examined a range of voltages and found that the average amplitude of the evoked EPSP steadily increases as a nearly linear

function of the somatic membrane potential of the presynaptic cell, ranging from 26 to 258% over the 15–24 mV range tested (Fig. 2a, $n = 5$). The average slope of the linear fits was $31.2 \pm 44.8\%$ enhancement of EPSP amplitude per 10 mV of presynaptic depolarization. EPSPs of all amplitudes showed facilitation (Fig. 2a), but there was no significant relationship between EPSP amplitude and per cent facilitation ($r = -0.22$, $P = 0.26$; data not shown). Examination of the distribution of amplitudes of single evoked EPSPs revealed that presynaptic-depolarization-induced facilitation resulted in a rightward shift of the EPSP amplitude histogram (Fig. 2b, c; $n = 15$) and cumulative EPSP amplitude distribution (Supplementary Fig. 3).

One possible mechanism for this facilitatory effect is that depolarization of the somata of pyramidal neurons results in depolarization of their nearby synaptic terminals. Depolarization is known to strongly increase transmitter release at central synapses (with an increase of 10% per mV in one study³), perhaps through small increases in intracellular calcium ($[\text{Ca}^{2+}]_i$). To test this possibility, we recorded from neurons using pipettes that contained the calcium chelators BAPTA (25 μM) and either 10 mM EGTA ($n = 16$) or 1 mM EGTA ($n = 11$). The concentrations of EGTA and BAPTA at the relevant sites of synaptic transmission between the two recorded neurons are not known and may be substantially lower. In cells recorded with 10 mM EGTA, somatic-depolarization-induced enhancement of EPSPs was observed in only 1/16 pairs ($P < 0.01$ by χ^2 test, compared to BAPTA alone), whereas facilitation was observed in 3/11 pairs recorded with 1 mM EGTA ($P = 0.08$).

Facilitation of action-potential-triggered PSPs by somatic depolarization has been observed in several invertebrate preparations and seems to occur through an increase in the probability of transmitter release, either by depolarization of synapses that are electrotonically close to the soma, or by broadening of the action potential^{4–7}. To investigate which of these may account for our facilitation effect, we obtained simultaneous $G\Omega$ seal, whole-cell recordings from the somata and axons of layer 5 pyramidal cells (for example, see Fig. 4c; $n = 76$). Axonal whole-cell recordings were possible because cortical axons form spherical 4–6- μm diameter ‘bleb’-like structures at the surface of the slice in response to the slicing procedure. Although this structure is certainly abnormal in some regards, whole-cell recordings from the axon bleb will reflect the action potential properties of the main axon, owing to the electrotonic compactness of these structures. Depolarization of the soma of layer 5 pyramidal neurons significantly increased the duration, decreased the amplitude and increased the integrated area ($\text{mV} \times \text{ms}$) of action potentials recorded in the axon, even up to 300 μm from the soma (Fig. 3), especially near the firing threshold. Changes in action potential duration with depolarization of the soma occurred very rapidly in the cell body, being complete within 100 ms (Fig. 3e, $n = 9$; see also Supplementary Fig. 2). In contrast, changes in action potential duration in the distal (90–400 μm) axon occurred more

¹Department of Neurobiology, Kavli Institute for Neuroscience, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510, USA.

slowly. Action potential broadening had a time constant of 2.3 s with depolarization of the soma, whereas hyperpolarization of the soma shortened axonal action potentials, also with a slow time course ($\tau = 1.5$ s, $n = 9$; Fig. 3f), even though changes in somatic membrane potential were rapidly (within 25 ms) reflected in the axonal membrane potential. Changes in the duration of axonal action potentials occurred through changes in the falling phase, and are suggestive of the slow inactivation and removal of inactivation of K^+ currents important for axonal action potential repolarization⁸.

It has previously been shown in cultured hippocampal pyramidal neurons that the A-current can gate axon conduction, resulting in apparent axon conduction failures from hyperpolarized membrane potentials⁹. However, we failed to observe any failures of orthodromic axon conduction along the main axon—even several hundred micrometres from the soma—at any of the membrane potentials tested (-80 to -50 mV, $n = 76$ paired soma/axon recordings; see also refs 10, 11). In addition, involvement of an A-current is not consistent with the time course of the enhancement effect we observed (Fig. 1f), nor with the rightward shift in the distribution of evoked EPSP amplitudes (Fig. 2b, c). Similarly, we did not find a significant effect of hyperpolarizing current pulses (100 ms, 15–20 mV) ending 10 ms before the onset of the 5-ms depolarizing current pulse that initiated action potentials in the presynaptic neuron ($n = 6$ pairs).

We next examined the possibility that naturally occurring barrages of synaptic potentials might have a significant effect on the membrane potential of distal axons. By reducing the bath concentration of Ca^{2+} and Mg^{2+} to 1 mM each, we were able to induce a slow oscillation of network activity known as Up and Down states^{12–15} ($n = 6$; Fig. 4a). This slow oscillation is associated with periods of recurrent excitatory and inhibitory activity within cortical networks,

interspersed with periods of relative quiescence, and occurs naturally during periods of slow-wave sleep¹⁶. *In vivo*, these Up and Down states range between 10 and 20 mV in amplitude^{15–17}, but in our submerged cortical slices they were of more modest magnitude (9.2 ± 3.2 mV, $n = 20$ cells), presumably owing to the extensive loss of intracortical connectivity.

Performing simultaneous axonal and somatic whole-cell recordings from layer 5 cortical pyramidal cells ($n = 18$) revealed that the voltage fluctuations associated with synaptic barrages propagated long distances (>0.4 mm) along the main axon (Fig. 4e). The axonal form of synaptic barrages was a close copy of that recorded in the soma, with some attenuation at higher frequencies in the axon (Fig. 4b, $n = 18$). The axonal:somatic ratio of the integrated amplitude (mV $\times$ ms) of Up state synaptic activity decreased with distance from the soma, and this decrease could be well fitted ($r^2 = 0.82$, $P < 0.01$) with an exponential function with a length constant (λ) of 417 μ m (Fig. 4e, $n = 18$), which is similar to the axonal length constant associated with constant depolarization of the soma (455 μ m, $n = 28$). Our computational simulations indicate that even cortical neurons with extensive axonal arbors may also have axonal length constants in the hundreds of micrometres for low frequencies (see Supplementary Fig. 4), indicating that these long length constants are not an artefact of axotomy during slicing.

To further examine the electrophysiological communication of activity from the soma to the axon, we used the dynamic clamp technique^{14,18} to inject Up-like conductance states (as well as broadband conductance noise) into the soma, and recorded the subsequent transfer of this activity down the axon (see Supplementary Figs 5 and 6; $n = 23$). These recordings revealed the rapid propagation of somatic membrane potential down the axon, again with attenuation at higher frequencies, with a decline in the overall amplitude time

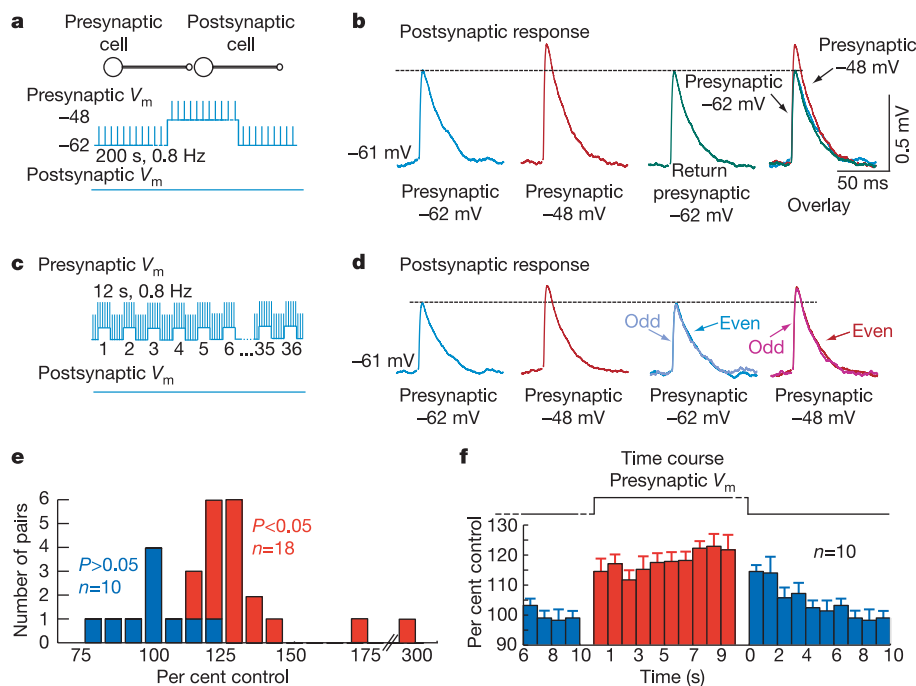


Figure 1 | Somatic depolarization results in an increase in the amplitude of evoked EPSPs in nearby neurons. **a**, A presynaptic layer 5 neuron was depolarized from rest (-62 mV in this cell) to near firing threshold (-48 mV in this cell), while being caused to spike at a rate of 0.8 Hz by the injection of 1-ms depolarizing current pulses. The resting and depolarized periods were 100–200 s each. **b**, Average EPSP amplitude at the two presynaptic somatic membrane potentials. See Supplementary Fig. 1 for morphology. **c**, The facilitating effect was also examined by intermixing 10–20-s periods of depolarization and hyperpolarization (typically 10–15 mV difference) of the

soma of the presynaptic cell. **d**, Depolarization again resulted in a significant increase in the amplitude of the average evoked EPSP. Comparing the average EPSPs evoked on odd and even blocks within each depolarized or hyperpolarized state does not reveal any within-state difference, illustrating the stability of this effect. **e**, The median facilitated amplitude was 129% of the control, ranged from 112 to 294% and was seen in 18/28 pairs. **f**, Average time course of the facilitation and disfacilitation of EPSPs in cells that underwent the protocol in **c** and showed significant enhancement (1-s bin widths are shown with s.e.m.; see Supplementary Fig. 2).

course of synaptic-like barrages that could be well fitted by an exponential function ($\lambda = 455 \mu\text{m}$). Calculating the frequency-dependent transfer ratio from somatic conductance to axonal membrane potential revealed a strong transfer of moderate to low ($<20 \text{ Hz}$) frequencies, and decreasing transfer with progressively higher frequencies (Supplementary Fig. 6; $n = 23$).

Layer 5 pyramidal neurons, like other cortical neurons, give rise to a local high density of axonal connections to other pyramidal and non-pyramidal cells^{19–23} (Supplementary Figs 1, 4 and 7). Indeed, examination of the main axon and local axon collaterals of biocytin-filled layer 5 pyramidal cells from our slices revealed, on average, 155 (± 79 ; $n = 14$ cells) putative synaptic boutons within 0.5-mm of the cell body, and 269 (± 152) putative synaptic boutons within the first 1 mm (Fig. 4c and Supplementary Fig. 7). These values are probably a significant underestimate of local synaptic connectivity, owing to the cutting of axons and limitations of axonal staining using the slice technique (see Supplementary Figs 1 and 4).

Traditionally, somatic-to-axonal synaptic terminal communication in the mammalian brain is thought to occur solely through the rate and timing of action potentials. Here we demonstrate a new and important role for presynaptic somatic membrane potential in synaptic communication between nearby cortical neurons. The long length constant of layer 5 pyramidal cell axons allows for the rapid dissemination of changes in membrane potential to substantial numbers of nearby synaptic terminals, as well as influencing the duration of axonal action potentials. We hypothesize that one or both of these alterations underlies the depolarization-induced enhancement of synaptic transmission observed here^{3,24,25}.

We propose that the ability of the presynaptic somatic membrane potential to affect the magnitude of synaptic potentials is not limited

to cortical layer 5 pyramidal cells, but may occur at many brain cell types that have presynaptic terminals that are electrotonically close to the cell body. This is especially relevant for the neocortex, where many different types of excitatory and inhibitory neurons form dense local synaptic networks^{19–23}. The ability of the presynaptic dendrosomatic membrane potential to significantly influence synaptic transmission indicates that cells function in a more 'holistic' manner than previously appreciated, and has important functional consequences. For example, cortical neuronal networks that enter periods of persistent activity through recurrent excitation/inhibition¹³—which might underlie working memory, premotor planning and decision making—may do so in part through a conjoint increase in synaptic amplitudes owing to depolarization of the network. In addition, the large membrane potential changes that occur in central neurons in response to state changes (for example, waking versus slow-wave sleep)¹⁶, the release of neuromodulatory neurotransmitters²⁶, sensory stimulation^{27,28}, epileptic seizures (see Supplementary Fig. 8) or spreading depression (a leading model for the aura of migraine headaches²⁹) may strongly and directly influence synaptic transmission between affected neurons. On the basis of our findings, we suggest that local interneuronal communication in the brain is a

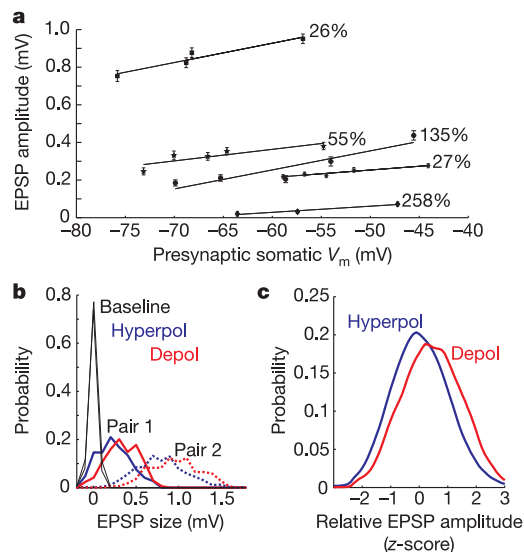


Figure 2 | Properties of EPSP enhancement. **a**, Plotting the average amplitude of the evoked EPSP versus the somatic membrane potential of the presynaptic neuron reveals continuous functions that are well fitted with a linear function ($n = 5$ pairs). To the right of each trace is the per cent enhancement of the EPSP recorded at the most depolarized, versus the most hyperpolarized, membrane potentials tested. Points show mean $\pm$ s.e.m. **b**, Two examples of histogram EPSP amplitude distributions at hyperpolarized and depolarized membrane potentials (following the protocol in Fig. 1a; both pairs show a significant shift, $P < 0.01$). The black line illustrates the level of baseline activity, using our measure of synaptic amplitude (see Methods). **c**, Depolarization significantly ($P < 0.01$) shifts the normalized EPSP amplitude histogram ($n = 15$ pairs) to the right. Normalization was performed by fitting the hyperpolarized amplitude distribution of each cell with a gaussian curve and then plotting the data such that the mean of the hyperpolarized state was 0 and the standard deviation was 1.

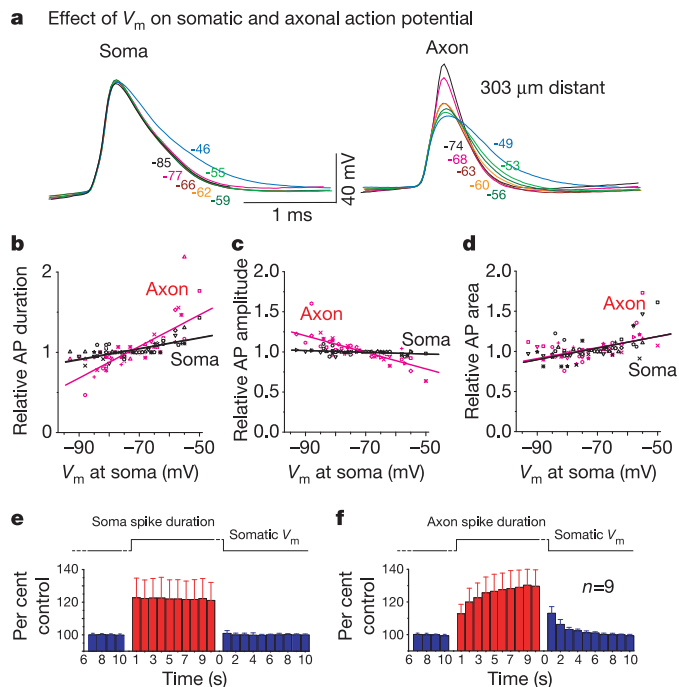


Figure 3 | Changes in somatic membrane potential affect the amplitude and duration of somatic and axonal action potentials. **a**, Simultaneous somatic and axonal (303 μm from soma) recording while the soma is depolarized or hyperpolarized to different levels by the injection of current. Action potentials are initiated with brief (5 ms) depolarizing current pulses through the somatic recording electrode. Raw, un-normalized traces are shown after at least 10 s with the baseline membrane potential at the value indicated. **b**, Plot of the duration of somatic (black) and axonal (red) action potentials, normalized to the duration recorded near -70 mV (soma $0.60 \pm 0.11 \text{ ms}$, axon $0.55 \pm 0.24 \text{ ms}$; measured at half height), versus the voltage at the soma for ten cells. **c**, Plot of the amplitude of somatic and axonal action potentials after normalization to the value obtained at a membrane potential of -70 mV (soma $91.2 \pm 6.3 \text{ mV}$; axon $71.0 \pm 14.4 \text{ mV}$) versus somatic membrane potential. **d**, Plot of the action potential area ($\text{mV} \times \text{ms}$) versus somatic membrane potential after normalization to the value obtained at -70 mV (soma $26.3 \pm 3.2 \text{ mV ms}^{-1}$; axon $16.4 \pm 3.9 \text{ mV ms}^{-1}$). **e**, **f**, Time course of changes in the duration of action potentials in the soma (**e**) and axon (**f**; average distance of 215 μm) during the indicated changes in membrane potential (see Supplementary Fig. 2). The average spike duration corresponding to 100% was 0.61 ms for soma, 0.54 ms for axon.

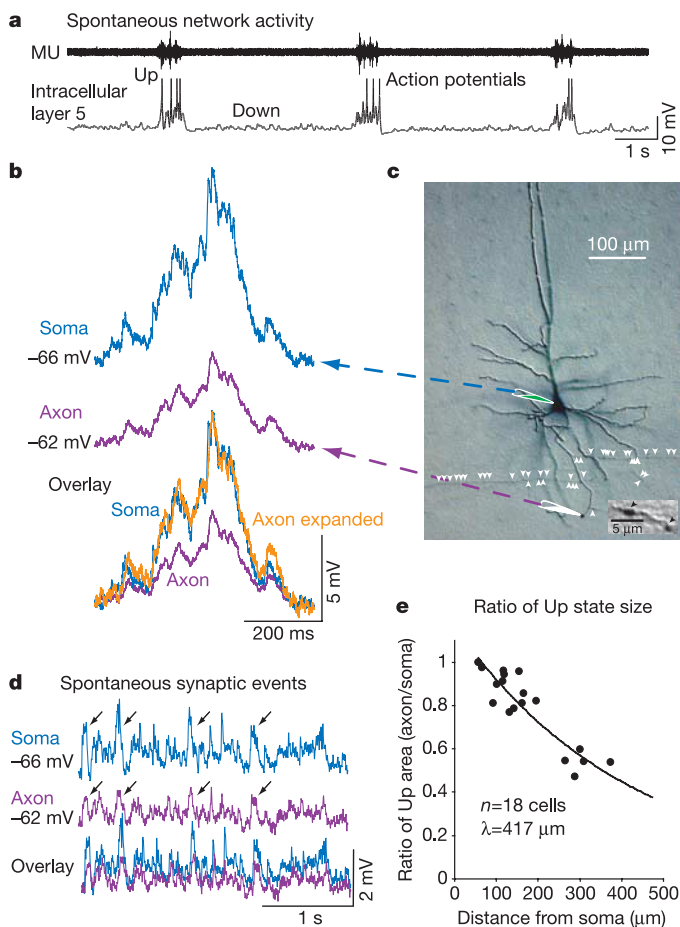


Figure 4 | Spontaneous barrages of synaptic activity propagate down the axon. **a**, Simultaneous extracellular multiple unit (MU) and somatic whole-cell recording in layer 5 during the generation of recurrent network activity of the Up state, interspersed with relatively quiet periods of the Down state. **b**, Simultaneous recording of an Up state in the soma and cut end of the main axon (266 μ m from the soma). The bottom traces are an overlay of axonal (purple) and somatic (blue) recordings of synaptic activity, along with a version of the axonal recording that has been expanded to match the peak amplitude of the somatic recording (orange). **c**, Morphology of the recorded cell. White arrowheads indicate putative presynaptic boutons along the axon collaterals in this focal plane. Inset shows an example of two putative *en passant* boutons. Note the large number of boutons that seem to be electronically close to the soma. **d**, Even during quiescent periods, spontaneous synaptic events (arrows point to examples) are recorded in both the soma and the cut end of the main axon of the cell in **c**. **e**, Comparison of the amplitude (area) of Up states at the soma and the cut end of the axon in different layer 5 neurons reveals an orderly decrease in the amplitude of spontaneous synaptic activity with distance from the soma. The calculated length constant of the axon is 417 μ m.

result of discrete, action-potential-dependent triggered release that is graded by presynaptic membrane potential.

Note added in proof: A recent study has demonstrated a similar phenomenon to that described here at the synaptic connection between granule cells and CA3 pyramidal cells in the hippocampus³¹. Together, these findings reveal the mixed triggered and graded (analogue) nature of synaptic transmission in the mammalian brain, even over relatively long distances.

METHODS

See Supplementary Information for full descriptions of the methods used. Experiments were performed using standard slice and patch-clamp procedures. Briefly, 0.3-mm thick slices of ferret cortex (7–10-weeks old), typically from the prefrontal region but also from the somatosensory area, were incubated at 35 °C

and then maintained *in vitro* in a submerged-style recording chamber at 36.5 °C on an upright infrared-differential interference contrast (IR-DIC) microscope equipped with fluorescence for the visualization of axonal profiles. The artificial cerebrospinal fluid (ACSF) contained 126 mM NaCl, 2.5 mM KCl, 2 mM MgSO₄, 2 mM CaCl₂, 26 mM NaHCO₃, 1.25 mM NaH₂PO₄, 25 mM dextrose (315 mOsm, pH 7.4). For only those recordings in which the slow oscillation was examined in the submerged chamber, the ACSF solution was modified to contain 1 mM MgSO₄, 1 mM CaCl₂ and 3.5 mM KCl^{12–14}. The membrane potentials in our whole-cell recordings were not corrected for Donnan liquid junction potentials, which typically range between 5 and 15 mV (ref. 30).

Whole-cell recordings were achieved from both soma and the cut end of the main axon using a Multiclamp 700B or Axoclamp 2B amplifier (Axon Instruments). Patch pipettes (somatic 5–6 M Ω , axonal 9–15 M Ω) were filled with 140 mM potassium gluconate, 3 mM KCl, 2 mM MgCl₂, 2 mM Na₂ATP, 10 mM HEPES, 0.025 mM BAPTA, Alexa Fluor 488 (100 μ M; for axonal recording experiments only), 0.2% biocytin, pH 7.2 with KOH (288 mOsm). Activation of nearby (within 60 μ m) morphologically identified pyramidal cells resulted in EPSPs that (1) were blocked by bath application of the non-NMDA glutamatergic receptor antagonist CNQX (10 μ M in bath; $n=4$), (2) decayed with an exponential time course, and (3) were not followed by inhibitory synaptic potentials (Fig. 1b). Data were analysed only if the evoked EPSP did not exhibit significant run-down or other instability, as shown by a statistically significant change in the amplitude of the EPSP either between the two control periods (when the presynaptic membrane potential was at rest before and after the depolarizing period) or between the first and last third of the interspersed control periods (as in the protocol of Fig. 1c).

For simultaneous somatic and axonal whole-cell recordings, the terminal 'bleb' of the axon on the surface of the slice was identified through brief (less than 20 s) exposure of the cell to fluorescence. Recordings with access resistance (monitored frequently) higher than 25 M Ω for somatic recording or 45 M Ω for axonal recording were discarded. Bridge balance and capacitance neutralization were carefully adjusted before and after every experimental protocol. After a recording was completed, the slice was transferred to 4% paraformaldehyde in 0.1 M phosphate buffer for subsequent immunostaining and visualization.

EPSP magnitudes were measured by first carefully identifying presynaptic action potential time courses. The average EPSP in the postsynaptic cell was calculated synchronized to the timing of presynaptic action potentials, and the timing of the peak of the average EPSP was determined. The amplitude of each evoked EPSP on single trials was taken as the difference between the postsynaptic membrane potential at the peak time of the average EPSP after the action potential and the membrane potential before onset of the current pulse evoking the action potential. Baseline activity was measured as the difference in membrane potential over the same time delay, but without activation of a presynaptic spike. Unless otherwise stated, statistical tests were performed using Student's *t*-tests. Values are presented as mean $\pm$ s.d. in the text and mean $\pm$ s.e.m. in the figures.

Received 30 December 2005; accepted 16 March 2006.

Published online 12 April 2006.

1. Stuart, G., Schiller, J. & Sakmann, B. Action potential initiation and propagation in rat neocortical pyramidal neurons. *J. Physiol. (Lond.)* 505, 617–632 (1997).
2. Mainen, Z. F., Joerges, J., Huguenard, J. R. & Sejnowski, T. J. A model of spike initiation in neocortical pyramidal neurons. *Neuron* 15, 1427–1439 (1995).
3. Awatramani, G. B., Price, G. D. & Trussell, L. O. Modulation of transmitter release by presynaptic resting potential and background calcium levels. *Neuron* 48, 109–121 (2005).
4. Nicholls, J. & Wallace, B. G. Quantal analysis of transmitter release at an inhibitory synapse in the central nervous system of the leech. *J. Physiol. (Lond.)* 281, 171–185 (1978).
5. Shimahara, T. & Tauc, L. Multiple interneuronal afferents to the giant cells in *Aplysia*. *J. Physiol. (Lond.)* 247, 299–319 (1975).
6. Shapiro, E., Castellucci, V. F. & Kandel, E. R. Presynaptic membrane potential affects transmitter release in an identified neuron in *Aplysia* by modulating the Ca²⁺ and K⁺ currents. *Proc. Natl Acad. Sci. USA* 77, 629–633 (1980).
7. Ivanov, A. I. & Calabrese, R. L. Modulation of spike-mediated synaptic transmission by presynaptic background Ca²⁺ in leech heart interneurons. *J. Neurosci.* 23, 1206–1218 (2003).
8. Storm, J. F. Temporal integration by a slowly inactivating K⁺ current in hippocampal neurons. *Nature* 336, 379–381 (1988).
9. Debanne, D., Guerinéau, N. C., Gähwiler, B. H. & Thompson, S. M. Action-potential propagation gated by an axonal I_A-like K⁺ conductance in hippocampus. *Nature* 389, 286–289 (1997).
10. Cox, C. L., Denk, W., Tank, D. W. & Svoboda, K. Action potentials reliably invade axonal arbors of rat neocortical neurons. *Proc. Natl Acad. Sci. USA* 97, 9724–9728 (2000).

11. Koester, H. J. & Sakmann, B. Calcium dynamics associated with action potentials in single nerve terminals of pyramidal cells in layer 2/3 of the young rat neocortex. *J. Physiol. (Lond.)* **529**, 625–646 (2000).
12. Sanchez-Vives, M. V. & McCormick, D. A. Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nature Neurosci.* **3**, 1027–1034 (2000).
13. Shu, Y., Hasenstaub, A. & McCormick, D. A. Turning on and off recurrent balanced cortical activity. *Nature* **423**, 288–293 (2003).
14. Shu, Y., Hasenstaub, A., Badoual, M., Bal, T. & McCormick, D. A. Barrages of synaptic activity control the gain and sensitivity of cortical neurons. *J. Neurosci.* **23**, 10388–10401 (2003).
15. Hasenstaub, A. *et al.* Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron* **47**, 423–435 (2005).
16. Steriade, M., Timofeev, I. & Grenier, F. Natural waking and sleep states: a view from inside neocortical neurons. *J. Neurophysiol.* **85**, 1969–1985 (2001).
17. Steriade, M., Nunez, A. & Amzica, F. A novel slow (<1 Hz) oscillation of neocortical neurons *in vivo*: depolarizing and hyperpolarizing components. *J. Neurosci.* **13**, 3252–3265 (1993).
18. Dorval, A. D., Christini, D. J. & White, J. A. Real-time linux dynamic clamp: a fast and flexible way to construct virtual ion channels in living cells. *Ann. Biomed. Eng.* **29**, 897–907 (2001).
19. Binzegger, T., Douglas, R. J. & Martin, K. A. Axons in cat visual cortex are topologically self-similar. *Cereb. Cortex* **15**, 152–165 (2005).
20. Binzegger, T., Douglas, R. J. & Martin, K. A. A quantitative map of the circuit of cat primary visual cortex. *J. Neurosci.* **24**, 8441–8453 (2004).
21. Gilbert, C. D. & Wiesel, T. N. Clustered intrinsic connections in cat visual cortex. *J. Neurosci.* **3**, 1116–1133 (1983).
22. Markram, H., Lubke, J., Frotscher, M., Roth, A. & Sakmann, B. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol. (Lond.)* **500**, 409–440 (1997).
23. Peters, A. & Jones, E. G. (eds) *Cerebral Cortex Vol. 1* (Plenum Press, New York, 1984).
24. Jackson, M. B., Konnerth, A. & Augustine, G. J. Action potential broadening and frequency-dependent facilitation of calcium signals in pituitary nerve terminals. *Proc. Natl Acad. Sci. USA* **88**, 380–384 (1991).
25. Geiger, J. R. & Jonas, P. Dynamic control of presynaptic Ca^{2+} inflow by fast-inactivating K^+ channels in hippocampal mossy fiber boutons. *Neuron* **28**, 927–939 (2000).
26. McCormick, D. A. Neurotransmitter actions in the thalamus and cerebral cortex and their role in neuromodulation of thalamocortical activity. *Prog. Neurobiol.* **39**, 337–388 (1992).
27. Carandini, M. & Ferster, D. A tonic hyperpolarization underlying contrast adaptation in cat visual cortex. *Science* **276**, 949–952 (1997).
28. Sanchez-Vives, M. V., Nowak, L. G. & McCormick, D. A. Membrane mechanisms underlying contrast adaptation in cat area 17 *in vivo*. *J. Neurosci.* **20**, 4267–4285 (2000).
29. Welch, K. M. Contemporary concepts of migraine pathogenesis. *Neurology* **61**, S2–S8 (2003).
30. Fricker, D., Verheugen, J. A. & Miles, R. Cell-attached measurements of the firing threshold of rat hippocampal neurones. *J. Physiol. (Lond.)* **517**, 791–804 (1999).
31. Alle, H. & Geiger, J. R. P. Combined analog and action potential coding in hippocampal mossy fibers. *Science* **311**, 1290–1293 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the NIH (D.A.M.), the Howard Hughes Foundation (A.H.) and the Kavli Institute for Neuroscience.

Author Contributions Y.S. performed all recordings, A.H. analysed the data, A.D. did cell reconstructions, Y.Y. performed computational models, and D.A.M. wrote the paper and helped design experiments. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.A.M. (david.mccormick@yale.edu).

LETTERS

Germline transmission of genetically modified primordial germ cells

Marie-Cecile van de Lavoie¹, Jennifer H. Diamond¹, Philip A. Leighton¹, Christine Mather-Love¹, Babette S. Heyer¹, Renee Bradshaw¹, Allyn Kerchner¹, Lisa T. Hooi¹, Terri M. Gessaro², Susan E. Swanberg², Mary E. Delany² & Robert J. Etches¹

Primordial germ cells (PGCs) are the precursors of sperm and eggs¹. In most animals, segregation of the germ line from the somatic lineages is one of the earliest events in development²; in avian embryos, PGCs are first identified in an extra-embryonic region, the germinal crescent, after approximately 18 h of incubation. After 50–55 h of development, PGCs migrate to the gonad and subsequently produce functional sperm and oocytes^{3,4}. So far, cultures of PGCs that remain restricted to the germ line have not been reported in any species^{5,6}. Here we show that chicken PGCs can be isolated, cultured and genetically modified while maintaining their commitment to the germ line. Furthermore, we show that chicken PGCs can be induced *in vitro* to differentiate into embryonic germ cells that contribute to somatic tissues. Retention of the commitment of PGCs to the germ line after extended periods in culture and after genetic modification combined with their capacity to acquire somatic competence *in vitro* provides a new model for developmental biology. The utility of the model is enhanced by the accessibility of the avian embryo, which facilitates access to the earliest stages of development and supplies a facile route for the reintroduction of PGCs into the embryonic vasculature. In addition, these attributes create new opportunities to manipulate the genome of chickens for agricultural and pharmaceutical applications.

Blood containing PGCs was collected from stage 14–17 (H&H; nomenclature used in ref. 7) Barred Plymouth Rock (BPR) chicken embryos and cultured in knockout (KO)-DMEM medium conditioned on buffalo rat liver (BRL) cells (which are known to produce leukaemia inhibitory factor), containing stem cell factor (SCF; 6 ng ml⁻¹) and human recombinant fibroblast growth factor (FGF; 4 ng ml⁻¹). PGCs were grown on a feeder of either Sandoz inbred mouse-derived thioguanine-resistant and ouabain-resistant (STO) fibroblasts or BRL cells. Twelve cell lines determined to be male by the absence of the Xho repeat in the W chromosome⁸ (data not shown) and two female cell lines were derived from 114 individual embryos. Cells in all of the lines display a round morphology, remain unattached (Fig. 1a) and can be successfully cryopreserved using conventional techniques. After 217 days in culture, karyotype analysis of one cell line (PGC13) revealed that it was diploid with a male ZZ sex chromosome constitution. To determine whether the cultured PGCs maintain their germ cell characteristics we analysed the expression of the germline-specific genes chicken vasa homologue (CVH)⁹ and deleted in azoospermia-like (DAZL). Polymerase chain reaction with reverse transcription (RT-PCR) analysis of PGCs after 32, 143 and 197 days of culture showed expression of DAZL and CVH (Fig. 1b), whereas at 166 days of culture, western blot analysis showed production of the CVH protein in the cultured cells (data not shown). Expression of CVH and an ovomucin-like protein (OLP)

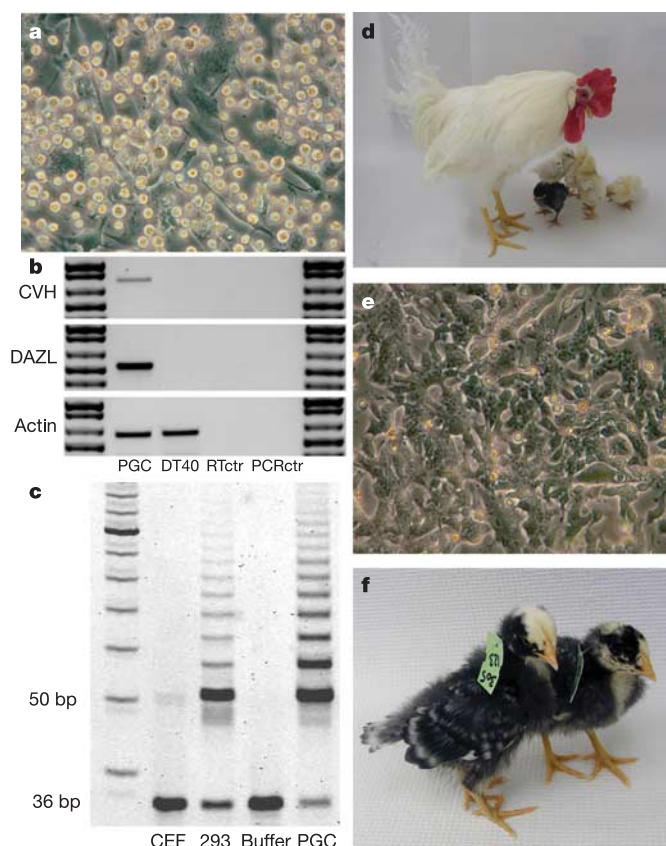


Figure 1 | Characteristics of PGC cell lines *in vitro*. **a**, PGCs are not attached, maintain a round morphology and are grown on a feeder layer of STO cells. **b**, CVH and DAZL expression in PGC13 and DT40 cells. Expression of both CVH and DAZL was observed in PGCs whereas DT40 cells did not express either CVH or DAZL. **c**, TRAP assay. Repeat sequences are visible in PGCs and the positive control cells (293), indicating the presence of telomerase. The 36-bp band in the negative control chicken embryonic fibroblast (CEF) lane is the internal PCR template. **d**, A WL is homozygous dominant at the dominant white locus (I/I). When bred to a BPR hen (i/i) all offspring from a WL rooster will be white (I/i). A black chick demonstrates that cultured PGCs derived from a BPR embryo (i/i) colonized the germ line. **e**, Chicken EG cells derived from PGCs in culture. EG cells are small, have large nuclei (light grey) and pronounced nucleoli. The EG cells were derived from PGCs isolated from black-feathered BPR embryos. **f**, Chimaeras obtained after EG cell injection into stage X (EG&K) white-feathered WL embryos. Somatic chimaerism is evident by the black feathers.

¹Origen Therapeutics, 1450 Rollins Road, Burlingame, California 94010, USA. ²Department of Animal Science, 1 Shields Avenue, University of California, Davis, 2131D Meyer Hall, Davis, California 95616, USA.

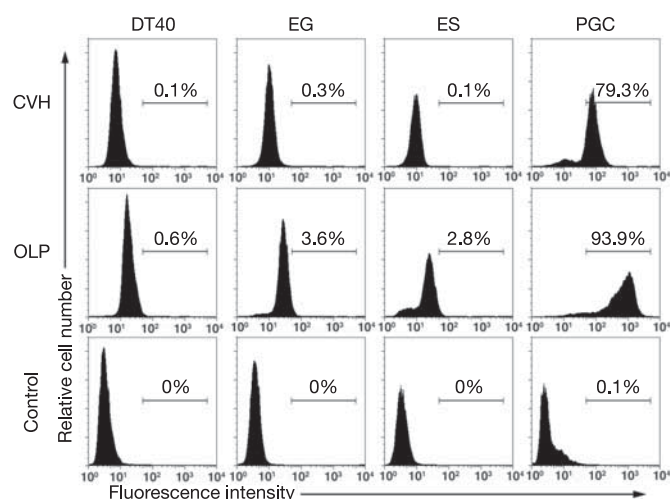


Figure 2 | FACS analysis of DT40 cells, ES cells, EG cells and PGCs, stained with antibodies against CVH and OLP. The DT40, ES and EG cells were negative for both markers, whereas a large majority of PGCs stained for both CVH and OLP.

that is expressed on the surface of PGCs during migration and colonization of the gonad¹⁰ were examined by fluorescence-activated cell sorting (FACS) at 44 and 65 days for PGC102, and at 229 and 280 days for PGC13. The FACS profiles of PGCs, chicken embryonic stem (ES) cells¹¹ and chicken embryonic germ (EG) cells revealed that the 1B3 antibody recognizes the ovomucin-like protein only on PGCs and that the expression of CVH is likewise restricted to PGCs (Fig. 2). By comparison, cells expressing an ES cell phenotype or an EG cell phenotype, and DT40 cells that express a pre-B-cell phenotype¹², did not express these antigens. Between 7% and 20% of the cells in the PGC cultures did not stain for either protein, indicating that some cells in each culture do not express a PGC phenotype. It is possible that the starting population was heterogeneous; however, a small part of the PGC population regularly attaches to the feeder layer and differentiates. This differentiation is also seen in clonally derived cell lines, indicating that the non-stained cells are differentiated PGCs. A telomeric repeat amplification protocol (TRAP) assay performed on PGCs that were in culture for 196 days demonstrated the presence of telomerase, which is a key characteristic of immortal cell lines (Fig. 1c). The presence of telomerase is consistent with the long-term growth of the PGCs in culture.

To demonstrate that PGCs maintain their restriction to the germ line, cells from eight male cell lines that were in culture for a minimum of 35 days and a maximum of 110 days were injected into stage 13–15 (H&H) White Leghorn (WL) embryos, which are post-gastrulation embryos with 19–27 somites and a functional vasculature. Twenty-four male chimaeric chicks were reared to sexual

maturity. Each of these chimaeras displayed the white-feathered phenotype of the WL breed (Fig. 1d). All of the male chimaeras transmitted the PGC phenotype (recognizable by black feathers) to the next generation, demonstrating that the cells retained their ability to colonize the germ line (Table 1 and Fig. 1d). The rate of germline transmission of these birds ranged from <1% to 86%, which is similar to the rates of germline transmission obtained in single-sex chimaeras after the transfer of freshly isolated PGCs¹³. Female PGCs in culture for 47 and 66 days were transmitted to the next generation by female chimaeras at frequencies up to 69% (Table 1), although these cell lines could not be maintained beyond 109 and 77 days, respectively. The variability of germline transmission among hens and roosters injected with aliquots of the same pool of cells may be due to variation in the number of injected PGCs that colonize the germ line and differential expansion of each of the cells within the gonad. Germline transmission of male PGCs has not been observed in 1,625 offspring of 14 female putative chimaeras, and female PGCs did not colonize the germ line of three male putative chimaeras that produced 2,739 offspring when mated to BPR hens. Similarly low rates of germline transmission were observed in mixed-sex chimaeras produced with freshly isolated PGCs¹³, indicating that the development of germ cells is impaired in mixed-sex chimaeras. To evaluate the reproductive capacity of offspring derived from cultured PGCs, male and female progeny were mated together. Fertility ranged from 53% to 85% and the hatch rate of fertile eggs ranged from 79% to 96%, indicating that the reproductive capacity of offspring derived from PGCs in culture for 40 days is normal.

Conventional electroporation protocols that we use routinely to introduce genetic modifications into chicken ES cells failed to yield genetically modified PGCs. To obviate potential silencing of transgenes, two copies of the HS4 insulator sequence from the chicken β -globin locus¹⁴ were inserted 5' and 3' of the transgenes and genetically modified clones were isolated. PGCs carrying an HS4- β -actin-EGFP- β -actin-puro-HS4 transgene (see Methods) were injected after 134 days in culture into stage 13–15 (H&H) embryos and eight roosters were bred to BPR hens to evaluate germline transmission. Seven of the eight roosters transmitted the PGC-derived genotype through the germ line at frequencies varying from 1% to 92%. Eighty of the 163 black-feathered offspring (49%) exhibited green fluorescence (Fig. 3b), indicating mendelian segregation of the transgene among PGC-derived offspring. Southern blot analysis confirmed that chicks expressing GFP were derived from the transgenic PGC line (Fig. 3a) and all of the chicks carrying a copy of the transgene expressed GFP (data not shown). We have also obtained transfected male cell lines carrying an HS4-ERNI-neo or an HS4- β -actin-puro transgene and have hatched live chicks carrying an HS4- β -actin-neo transgene from PGCs in culture for 267 days and an HS4-CAG-EGFP-CAG-puro transgene from PGCs in culture for 238 days.

A small proportion of PGCs attaches spontaneously to the feeder layer and assumes the morphology of ES cells, which is shown in

Table 1 | Germline transmission of primordial germ cells injected into the vasculature of stage 14–15 (H&H) embryos

Cell line	Sex	Time in culture (days)	No. of cells injected	No. of putative chimaeras tested	Germline transmission* (%)
PGC13	Male	40	1,200	3	0.1,1.5,17
PGC13	Male	110	2,500–3,000	5	1.1,1.5,3,84
PGC21	Male	44	1,500	3	10,16,21
PGC34	Male	47	3,000	3	42,74,80
PGC35	Male	35	3,000	7	15,23,47,61,80,85,86
PGC51	Male	47	3,000	1	11
PGC54	Male	47	3,000	4	0.5,2,20,24
PGC80	Male	29	3,000	1	55
PGC84	Male	50	3,000	1	70
PGC56	Female	66	3,000	5	1,3,6,52,69
PGC85	Female	47	3,000	10	0.0,1,1,1,4,5,10,11,12

*Each value represents the rate of germline transmission of one chimaera.

ref. 11. By removing FGF, SCF and chicken serum these cells can be expanded and are called EG cells (Fig. 1e). Without these components and with an increase in the percentage of conditioned medium, the culture conditions are the same as those used for long-term

culture of ES cells derived from blastodermal cells¹¹. Chicken EG cells are observed in both newly derived and clonally derived transgenic PGC lines. Southern blot analysis of EG cells derived from a clonal GFP-positive PGC line (Fig. 3a) indicates that EG cells originate from the PGCs and are not derived from a contaminating population of cells in blood samples from stage 14–17 (H&H) embryos from which the cultures originated. When EG cells were injected into stage X (EG&K; nomenclature used in ref. 15) embryos, which are the functional equivalents of mammalian blastocysts, they contributed extensively to various somatic tissues (Fig. 1f and Supplementary Fig. 1). Of the 140 WL embryos injected with EG cells, 30 survived until at least embryonic day (E)14, and 20 embryos (67%) exhibited the black feathering of the BPR breed that was used to derive the EG cells. Four of these chimaeras with extensive feather pigmentation have been bred to evaluate contributions to the germ line. One black chick was observed among 4,490 offspring, indicating that some capability to colonize the germ line is retained within the culture. In general, however, the ability of EG cells to colonize the somatic tissues and their inability to colonize the germ line is similar to that of chicken ES cells¹¹.

To determine whether cultured PGCs contribute to somatic tissues, tissues from chicks injected with GFP-positive PGCs were analysed by histology. With the exception of four green cells in a brain sample, cells expressing GFP were not present in other tissues, demonstrating that cells in a PGC culture remain restricted to the germ line. When PGC13 cells, after 209 days of culture, were injected into the subgerminal cavity of 111 stage X (EG&K) embryos, 20 embryos survived to E14 but black feathers were not observed, indicating that PGCs do not contribute to somatic tissues, even when injected at very early stages of development. However, three out of four roosters injected with PGCs at stage X (EG&K) transmitted them through the germ line at frequencies of 0.15%, 0.2% and 0.45%, indicating that they can colonize the germ line of early stage embryos. In contrast, chicken ES cells contribute substantially to somatic tissues but not to the germ line¹¹.

We describe a novel system for the production of transgenic chickens using PGCs. To our knowledge this is the first description in any species of PGCs that can be grown indefinitely *in vitro* and be genetically manipulated while retaining their commitment to the germ line. Combined with the extensive anatomical database describing development in the chick embryo and the ease of access to the earliest stages of development¹⁶, the ability to make genetic changes to chickens will provide a unique resource to address important issues in developmental biology. In addition, the ability to use cultured PGCs to make changes to the germ line of chickens has enormous implications for industrial¹⁷ and agricultural applications of avian transgenic technology.

METHODS

Derivation, culture and cryopreservation of PGCs. One to five microlitres of blood was taken from the vasculature of stage 14–17 (H&H) embryos and deposited in single wells of a 96- or 48-well plate seeded with either mitotically inactivated STO cells (3×10^4 cells cm^{-2}) or BRL cells (10^5 cells cm^{-2}). After 1–2 weeks, red blood cells had died and PGCs became visible. Throughout derivation and culture, the cells were grown in KO-DMEM (Invitrogen) that was conditioned with BRL cells¹⁸. The medium was supplemented with 7.5% fetal calf serum (FCS), 2.5% chicken serum, 2 mM glutamine, 1 mM pyruvate, 1× nucleosides, 1× non-essential amino acids and 0.1 mM β -mercaptoethanol, 6 ng ml^{-1} SCF and 4 ng ml^{-1} human recombinant FGF. For passage, the cells and medium were removed into a centrifuge tube, pelleted by centrifugation at 300 g for 5 min, resuspended and seeded at a concentration of 25,000 cells cm^{-2} . For cryopreservation, the cells were resuspended in CO₂-independent medium containing 10% FCS, 1.0% penicillin/streptomycin and 10% DMSO. The vials were frozen at -80°C and transferred to LN₂ after 24 h.

RNA isolation and RT-PCR. A 750-base-pair (bp) fragment from the CVH transcript, a 536-bp fragment from the DAZL transcript and a 597-bp fragment from the chicken β -actin transcript were amplified by RT-PCR (Supplementary Table 1).

FACS analysis of PGCs, EG cells and ES cells. Cells were washed in PBS/2% FBS

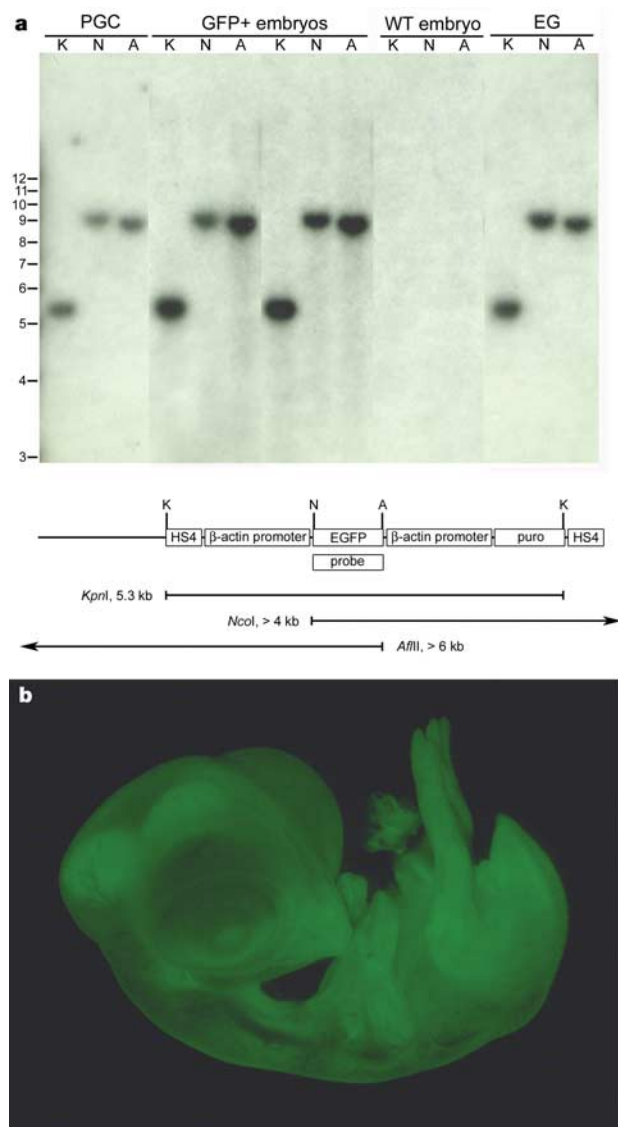


Figure 3 | β -Actin-EGFP is incorporated into the PGC genome and expressed in offspring. **a**, Southern blot analysis showing that a clonally derived, transfected PGC line can contribute to the germ line in chimaeric chickens and differentiate into EG cells. Top panel: samples of genomic DNA from PGCs transfected with the HS4- β -actin-EGFP- β -actin-puro construct, three embryos derived from a chimaeric rooster made with the transfected PGCs, and EG cells derived from the transfected PGCs were digested with restriction enzymes for detecting internal (*KpnI*) and junction fragments (*NcoI*, *AflII*) of the transgene insertion. The digested DNA was separated on a 0.7% agarose gel, blotted to nylon membrane, and probed with radiolabelled EGFP sequences. The sizes of the hybridizing fragments in kilobases (left axis) were identical in the PGCs, EG cells and two embryos that showed green fluorescence (GFP⁺ embryos). A third, non-fluorescing embryo (WT embryo) showed no hybridization. Bottom panel: a schematic of the construct is shown, with the locations of the restriction sites indicated, and the expected restriction fragment sizes shown below. **b**, The offspring of a chimaeric rooster transgenic for the HS4- β -actin-EGFP- β -actin-puro transgene at embryonic stage 34 (H&H). GFP is expressed ubiquitously throughout the embryo. The image was captured on a Leica MZ16FA microscope fitted with a Leica DFC300FX camera and a PlanAPO 0.6× lens that were provided by Leica.

and fixed in 4% paraformaldehyde for 5 min. Cell aliquots to be stained for CVH were permeabilized with 0.1% Triton X-100 for 1–2 min. Primary antibody was added for 20 min, cells were washed twice and incubated with secondary antibody (Alexa 488 anti-rabbit IgG for CVH and control and Alexa 488 anti-rabbit IgM for 1B3) for 15 min. Control cells were stained only with secondary antibody.

Telomerase detection. Primordial germ cells that were in culture for 196 days were pelleted and washed with PBS before being frozen at -80°C until analysis. Cell extracts were prepared according to the manufacturer's directions using the TRAPeze telomerase detection kit (Serologicals Corporation)¹⁹. The positive control was the transformed human kidney cell line 293 and the negative control contained lysis buffer and an internal PCR control template.

Karyotype analysis. PGCs were incubated in $0.1\text{ }\mu\text{g ml}^{-1}$ colcemid (Karyomax, Invitrogen) for 2 h. The cells were pelleted and resuspended in 0.56% KCl solution. After 25 min, the cells were centrifuged for 6 min at 200 g and the pellet was fixed in 3:1 methanol:glacial acetic acid. For metaphase analysis the cells were dropped on slides, stained with Giemsa and analysed for the four pairs of macro-chromosomes (GGA1–GGA4) and the sex chromosomes.

Transfection of PGCs. A total of 5×10^6 PGCs were resuspended in 400 μl electroporation buffer (Speciality Media) to which 20 μg of linearized DNA was added. One exponential decay pulse (200 V, with 900–1,100 μF) or eight square wave pulses (250–350 V, 100 μs) were given. After transfection the cells were grown for several days before neomycin (300 $\mu\text{g ml}^{-1}$) or puromycin (0.5 $\mu\text{g ml}^{-1}$) was added. In most cases, single colonies were derived by plating the cells at low concentration in 48-well plates after transfection.

Transgene construction. Transgenes were flanked by the insulator element from the chicken β -globin locus. The 250-bp core sequence of hypersensitive site 4 (HS4) was amplified from chicken genomic DNA by PCR (Supplementary Table 1) and a tandem duplication of the HS4 site was made by joining two copies of the PCR product, in the same orientation, resulting in 2 $\times$ HS4. The 2 $\times$ HS4 insulator was then inserted at both the 5' and 3' ends of the following constructs: β -actin-puro, β -actin-neo (gifts from J.-M. Buerstedde), β -actin-EGFP- β -actin-puro, ERNI-neo²⁰, CAG-EGFP-CAG-puro²¹.

Production of chimaeras using PGCs. PGCs were injected using a 37- μm diameter needle into the anterior portion of the sinus terminales of a stage 13–15 (H&H) embryo. The injected embryos were transferred to a second surrogate shell for incubation until hatching^{11,22}.

Derivation of EG cells and production of chimaeras. When an EG cell phenotype became visible, the growth factors and chicken serum were removed from the medium. Chicken EG cells were grown in 80% BRL conditioned medium on irradiated STO feeders plated at 10^4 cells cm^{-2} (ref. 11). At confluency the cells were incubated in Ca/Mg free PBS to obtain small clumps and passaged 1:2 or 1:3. For the production of chimaeras the cells were trypsinized into a single cell suspension and resuspended at 5,000 cells μl^{-1} . One microlitre was injected into the subgerminal cavity of an irradiated stage X (EG&K) embryo. After injection the embryo was incubated for 3 days before being transferred to a second surrogate shell and incubated until hatching²².

Received 30 September 2005; accepted 12 April 2006.

- Nieuwkoop, P. & Sutasurya, L. *Primordial Germ Cells in the Chordates* (Cambridge Univ. Press, New York, 1979).
- Wylie, C. Germ cells. *Cell* **96**, 165–174 (1999).
- Fujimoto, T., Ninomiya, T. & Ukeshima, A. Observations of the primordial germ cells in blood samples from the chick embryo. *Dev. Biol.* **49**, 278–282 (1976).
- Swift, C. H. Origin and early history of the primordial germ-cells in the chick. *Am. J. Anat.* **15**, 483–516 (1914).
- McLaren, A. The quest for immortality. *Nature* **359**, 482–483 (1992).
- McLaren, A. Primordial germ cells in the mouse. *Dev. Biol.* **262**, 1–15 (2003).
- Hamburger, V. & Hamilton, H. A series of normal stages in the development of the chick embryo. *J. Morphol.* **88**, 49–92 (1951).
- Kodama, H. et al. Nucleotide sequences and unusual electrophoretic behavior of the W chromosome-specific repeating DNA units of the domestic fowl, *Gallus gallus domesticus*. *Chromosoma* **96**, 18–25 (1987).
- Tsunekawa, N., Naito, M., Sakai, Y., Nishida, T. & Noce, T. Isolation of chicken vasa homolog gene and tracing the origin of primordial germ cells. *Development* **127**, 2741–2750 (2000).
- Halfter, W. et al. An ovomucin-like protein on the surface of migrating primordial germ cells of the chick and rat. *Development* **122**, 915–923 (1996).
- van de Lavoie, M.-C. et al. High grade transgenic somatic chimeras from chicken embryonic stem cells. *Mech. Dev.* **123**, 31–41 (2006).
- Baba, T. W., Giroir, B. P. & Humphries, E. H. Cell lines derived from avian lymphomas exhibit two distinct phenotypes. *Virology* **144**, 139–151 (1985).
- Naito, M. et al. Differentiation of donor primordial germ cells into functional gametes in the gonads of mixed-sex germline chimaeric chickens produced by transfer of primordial germ cells isolated from embryonic blood. *J. Reprod. Fertil.* **117**, 291–298 (1999).
- Burgess-Beusse, B. et al. The insulation of genes from external enhancers and silencing chromatin. *Proc. Natl Acad. Sci. USA* **99** (suppl. 4), 16433–16437 (2002).
- Eyal-Giladi, H. & Kochav, S. From cleavage to primitive streak formation: a complimentary normal table and a new look at the first stages of the development of the chick. *Dev. Biol.* **49**, 321–337 (1973).
- Stern, C. D. The chick: A great model system becomes even greater. *Dev. Cell* **8**, 9–17 (2005).
- Zhu, L. et al. Production of human monoclonal antibody in eggs of chimeric chickens. *Nature Biotechnol.* **23**, 1159–1169 (2005).
- Petitje, J. N. in *Handbook of Stem Cells* (eds Lanza, R. et al.) 471–477 (Elsevier Academic Press, Burlington, Massachusetts, 2004).
- Kim, N. et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011–2014 (1994).
- Streit, A., Berliner, A. J., Papanayotou, C., Sirulnik, A. & Stern, C. D. Initiation of neural induction by FGF signalling before gastrulation. *Nature* **406**, 74–78 (2000).
- Niwa, H., Yamamura, K. & Miyazaki, J. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108**, 193–199 (1991).
- Rowlett, K. & Simkiss, K. Explanted embryo culture: *in vitro* and *in ovo* techniques for domestic fowl. *Br. Poult. Sci.* **28**, 91–101 (1987).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Gitter for technical assistance with the karyotype; A. Pradas-Monne for help in the laboratory; W. Halfter for providing the 1B3 antibody; Leica for the provision of optical equipment to photograph the GFP-positive embryo; and J.-M. Buerstedde for supplying β -actin-neo and β -actin-puro. This work was supported by the Small Business Innovation Research Programs of the USDA and the NIH to Origen Therapeutics and a USDA grant to M.E.D.

Author Contributions M.C.L. developed the cell culture system with the assistance of J.H.D., P.A.L. and R.B.; C.M.-L. and J.H.D. performed the embryological manipulations; P.A.L. executed the molecular biology in collaboration with B.S.H. and L.T.H.; A.K. provided animal care; T.M.G., S.E.S. and M.E.D. conducted the telomerase assay and karyotyping; M.C.L. and R.J.E. coordinated the contributions of authors and wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.J.E. (ratches@origentherapeutics.com).

LETTERS

SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope

Ghislain G. Cabal^{1*}, Auguste Genovesio^{2*†}, Susana Rodriguez-Navarro^{4†}, Christophe Zimmer², Olivier Gadal¹, Annick Lesne⁵, Henri Buc³, Frank Feuerbach-Fournier¹, Jean-Christophe Olivo-Marin², Eduard C. Hurt⁴ & Ulf Nehrbass⁶

Changes in the transcriptional state of genes have been correlated with their repositioning within the nuclear space¹. Tethering reporter genes to the nuclear envelope alone can impose repression² and recent reports have shown that, after activation, certain genes can also be found closer to the nuclear periphery^{3–6}. The molecular mechanisms underlying these phenomena have remained elusive. Here, with the use of dynamic three-dimensional tracking of a single locus in live yeast (*Saccharomyces cerevisiae*) cells, we show that the activation of *GAL* genes (*GAL7*, *GAL10* and *GAL1*) leads to a confinement in dynamic motility. We demonstrate that the *GAL* locus is subject to sub-diffusive movement, which after activation can become constrained to a two-dimensional sliding motion along the nuclear envelope. RNA-fluorescence *in situ* hybridization analysis after activation reveals a higher transcriptional activity for the peripherally constrained *GAL* genes than for loci remaining intranuclear. This confinement was mediated by Sus1 and Ada2, members of the SAGA histone acetyltransferase complex, and Sac3, a messenger RNA export factor, physically linking the activated *GAL* genes to the nuclear-pore-complex component Nup1. Deleting *ADA2* or *NUPI* abrogates perinuclear *GAL* confinement without affecting *GAL1* transcription. Accordingly, transcriptional activation is necessary but not sufficient for the confinement of *GAL* genes at the nuclear periphery. The observed real-time dynamic mooring of active *GAL* genes to the inner side of the nuclear pore complex is in accordance with the 'gene gating' hypothesis⁷.

To analyse directly the possibility of activation-dependent changes in the dynamic motility of genes, and to gain a better understanding of activation-dependent nuclear repositioning at a molecular level, we developed a dedicated imaging approach (see Methods). In brief, by combining real-time confocal microscopy with advanced image analysis we were able to track genes within the three-dimensional (3D) space of the nuclear volume of live yeast cells. Because active *GAL* genes have previously been shown to appear closer to the nuclear periphery in fixed cells⁴, we adopted this *GAL* system as a model to explore both the dynamics and the molecular basis of the observed phenomenon. One array of repeated *Tet* operator (*TetO*) was therefore inserted downstream of the *GAL1* gene in a strain expressing the nuclear pore protein Nup49 fused to green fluorescent protein (GFP) (Fig. 1a). This array was revealed *in vivo* by expression of the *Tet* repressor protein coupled to GFP⁸ (Fig. 1a) and subjected to automatic 3D detection and localization over time with correction for Z-detection artefacts (Methods and Supplementary Fig. 1). Visual inspection of the trajectories of *GAL* genes over time revealed a

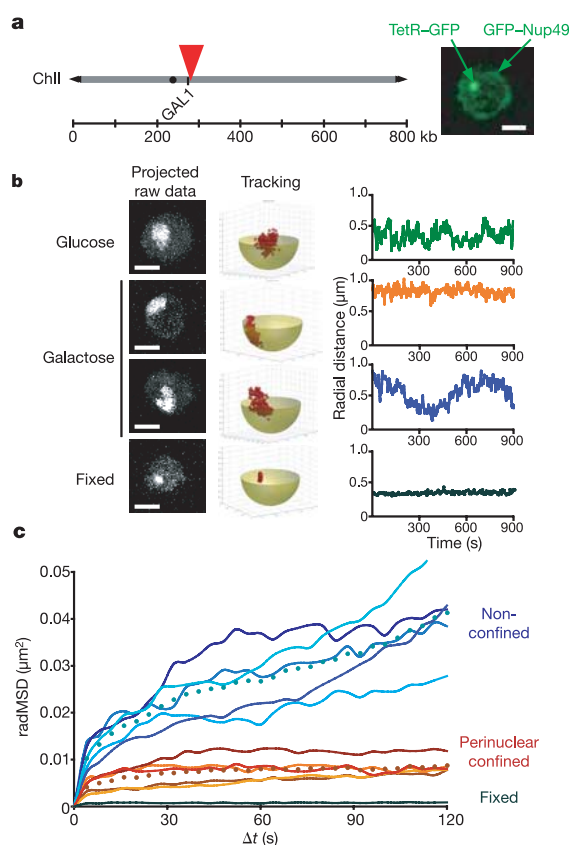


Figure 1 | *In vivo* dynamic analysis of movements of *GAL* genes in three dimensions. **a**, The *GAL* genes were tagged with TetO repeats (red arrowhead). This array labelled by TetR–GFP and the NE labelled with GFP–Nup49 were revealed simultaneously with a Nipkow-disk confocal microscope. kb, kilobases. Scale bar, 1 μ m. **b**, Z and time projection of the whole 3D raw data sequence (left), 3D plot of tracking sequence (centre), and distance between the centroid of the nucleus and the *GAL* genes plotted as a function of time (right) are shown for representative cells (strain yGC105). The entire set of tracking sequences is shown in Supplementary Fig. 2. Scale bar, 1 μ m. **c**, Individual radMSD (Supplementary Fig. 3) computed for each tracking sequence of *GAL* genes in galactose growth condition. The average for each indicated population is plotted as a dotted line.

¹Unité de Biologie Cellulaire du Noyau, ²Unité d'Analyse d'Images Quantitative, and ³Département de Biologie Cellulaire et Infection, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris cedex 15, France. ⁴Biochemie-Zentrum der Universität Heidelberg (BZH), Im Neuenheimer Feld 328, D-69120 Heidelberg, Germany. ⁵Laboratoire de Physique Théorique de la Matière Condensée, 4 place Jussieu, 75252 Paris cedex 05, France. ⁶Institut Pasteur Korea, 39-1, Hawolgok-dong, Sungbuk-gu, Seoul 136-791, Korea. [†]Present addresses: Institut Pasteur Korea, 39-1, Hawolgok-dong, Sungbuk-gu, Seoul 136-791, Korea (A.G.); Centro de Investigación Príncipe Felipe, Avda Salar 16, 46013 Valencia, Spain (S.R.-N.).

*These authors contributed equally to this work.

consistent and marked difference in the dynamic mobility between the repressed state (glucose growth condition) and the actively transcribed state (galactose growth condition, Supplementary Movies). Under glucose repression the tagged locus explored about one-third of the nuclear volume and was often found in the centre of the nucleus (Supplementary movie 1). By contrast, tracks recorded for *GAL* loci under galactose activation often had an exclusive perinuclear localization (Supplementary movie 2). The discrete trajectories of this type of movement usually followed along the nuclear envelope (NE), seemingly being barred from moving back into the nuclear interior. The dynamic motility of a gene can therefore become confined on activation. The type of tethering observed here is not static, because a spatially constrained movement continues along the NE. Rather, our data imply an activation-dependent, dynamic link between the transcribed genes and the NE, permitting a two-dimensional sliding movement underlying the envelope.

When tracking a larger number of cells over time we found that the dynamic change in the movement of activated *GAL* genes is not uniform. Although peripheral confinement was observed in half of the sampled cells, the dynamic motility in the remaining population turned out to be similar to that observed in glucose grown cells (Fig. 1b, and Supplementary Fig. 2). The mean squared change in radial distance (radMSD; Methods and Supplementary Fig. 3a)

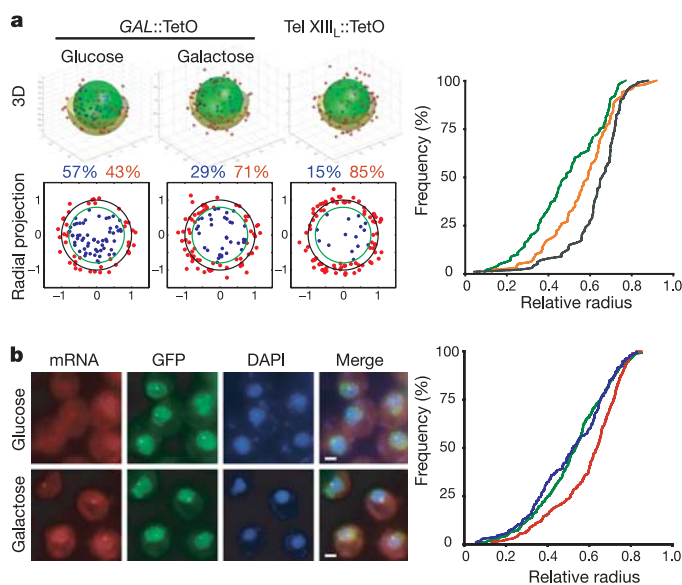


Figure 2 | Localization and expression of *GAL* genes on repression or activation. **a**, After Z-stack acquisitions, dedicated software was used to localize throughout the nuclear volume the labelled sub-telomere *Tel XIII_L* (strain yGC77) and the *GAL* genes (strain yGC75) in glucose-containing or galactose-containing medium in 100 wild-type cells. Left: positions of the tagged loci plotted in both 3D (top) and radial projections (bottom) (see Methods and Supplementary Fig. 1). The nucleus was divided into two equal volumes and the proportion of loci found in each volume is indicated. Right: cumulative distribution functions of the position of the *GAL* locus ($n = 100$; see Methods). Green, *GAL* in glucose-containing medium; orange, *GAL* in galactose-containing medium; black, sub-telomere *Tel XIII_L*. **b**, left: RNA-FISH on wild-type strain (yGC157) containing a galactose-dependent *lacZ* gene and the TetO array inserted at the *GAL* locus. *lacZ* mRNA (red), TetR-GFP and GFP-Nup49 (green) and DNA (blue; DAPI, 4,6-diamidino-2-phenylindole) are shown. Scale bar, 2 μ m. Right: cumulative distribution function of *GAL* locus position in glucose ($n = 385$), in galactose when mRNA is not detectable ($n = 178$) and in galactose when mRNA is visible ($n = 360$) are plotted. Pairwise comparisons of these distributions were computed with Kolmogorov-Smirnov tests: galactose with mRNA (red) versus galactose without mRNA (blue), $P = 8.75 \times 10^{-7}$; glucose (green) versus galactose without mRNA, $P = 0.2771$; glucose versus galactose with mRNA, $P = 1.23 \times 10^{-14}$.

plotted as a function of the time interval shows that the radial constraint of the movement does not occur in all cells under activating conditions (Fig. 1c).

To rule out the possibility of a detection artefact or bias inherent in the small sample analysis of dynamic studies, we performed static analyses permitting the statistical investigation of large numbers of cells. To verify the validity of our static approach we compared the extent of galactose-activation-dependent recruitment to the periphery with the previously reported perinuclear positioning of telomeres⁹. As expected, the left end of chromosome XIII was found in the peripheral volume in 85% of the analysed cells (Fig. 2a). Galactose-activated genes were found at the periphery in 71% of the nuclei, whereas repressed genes were evenly distributed throughout the nuclear volume, with only 43% of the loci at the periphery (Fig. 2a). Furthermore, the galactose-dependent recruitment seemed to be independent of the induction time (Supplementary Fig. 4). Taken together, these findings indicate that the activation-dependent recruitment of *GAL* genes might be transient and/or more labile than telomere anchoring.

The apparent inconsistency in the dynamic positioning of *GAL* genes within the same activated population of cells could imply one or both of the following: that genes spend a significant time not being transcribed even though they have been activated (as postulated previously¹⁰) or that only the peripheral loci are expressed. To address these assumptions directly, we performed RNA-fluorescence *in situ* hybridization (FISH) analyses with a probe specific to one single galactose-dependent construct inserted at the *GAL* locus. This experiment allowed us to observe that, on activation, the galactose-dependent transcripts are detected in only two-thirds of the cells (Fig. 2b). Furthermore, the transcribed genes are preferentially found in peripheral positions, whereas loci without detectable mRNA are randomly distributed, which is similar to results with cells grown in glucose-containing medium (Fig. 2b). Taken together, these results demonstrate that the transcriptional state is directly correlated with the position within the nucleus. Apparently, *GAL* genes are preferentially transcribed while their motility is being confined to the NE.

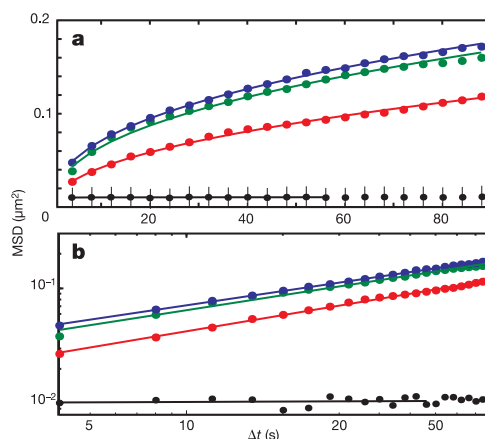


Figure 3 | Mean-squared-displacement analysis of movements of *GAL* genes. **a**, The MSDs (Supplementary Fig. 3) of the *GAL* genes averaged from ten cells grown in glucose (green), five cells grown in galactose for which the *GAL* genes were non-peripheral (blue), five cells grown in galactose for which the *GAL* genes stayed at the periphery (red), and one cell fixed with formaldehyde (black) are plotted in the same graph (strain yGC105). Fitted curves are computed for $\Delta t \leq 90$ s. The equations of the curve are as follows: glucose, $y = 0.024x^{0.43}$; non-confined galactose, $y = 0.027x^{0.41}$; perinuclear-confined galactose, $y = 0.014x^{0.47}$; fixed, $y = 0.01x^{0.01}$. **b**, The same set of data plotted with log-log axes. The entire 900-s sequence was used for the calculation but only Δt values from 0 to 90 s are shown.

To characterize the motility of *GAL* genes quantitatively and to enable a better comparison of active and inactive states, the mean squared displacements (MSDs; Methods and Supplementary Fig. 3b) were computed for loci in a repressed state, for loci in an activated state that appeared 'perinuclearly confined', and for 'non-confined' activated loci (Fig. 3). Previous studies have described the dynamics of genetic loci as confined random walks showing a free diffusion motion within the region of confinement¹. In this case, the expected MSD increases linearly with Δt at small time scales (namely $\langle \Delta d^2 \rangle = c(\Delta t)$ with a coefficient of diffusion $D = c/6$) and reaches a plateau at longer times¹¹. By contrast, our experimental data show a markedly different behaviour, with MSD curves at small time scales (less than 90 s) being well fitted by a power-law (namely $\langle \Delta d^2 \rangle = c(\Delta t)^\alpha$) with $\alpha \approx 0.4$, indicating sub-diffusion^{12,13} (Fig. 3). This sub-diffusion movement is observed in all three classes, showing that it depends neither on the transcriptional state nor on the perinuclear position of the *GAL* genes. The difference in motility between the two galactose classes is mostly reflected by a change of the prefactor (c). It decreases to roughly two-thirds for *GAL* genes classified as peripheral, which is consistent with the motion's being restricted to an area along the NE (two-dimensional) instead of a portion of the nuclear volume (3D). In consequence, our data show that both active and inactive *GAL* genes are subject to constrained movement, which cannot be described as a 'confined random walk'. Moreover, spatial repositioning did not at all affect the motility rate of the activated gene. Rather, the active gene sustained its sub-diffusive motility rate in a two-dimensional plane unabated by the presumed perinuclear tether.

The observed repositioning and dynamic confinement of *GAL* loci are clearly linked with transcriptional activation. Activation of *GAL* genes is mediated by the SAGA histone acetyltransferase complex^{14–16}. The Sus1 protein was found as a SAGA member interacting directly with the Sac3–Thp1–Cdc31 mRNA export complex, which is

bound to the nuclear pore complex (NPC) through Nup1 (refs 17–19). Furthermore, Sus1 is required for full *GAL1* gene expression and is physically present at the *GAL1*–*GAL10* promoter after transcriptional activation¹⁹. To study the effect of those interactions on the positioning of *GAL* genes, we analysed the 3D localization of *GAL* loci in *sac3Δ* and *sus1Δ* mutant strains. We found the galactose-dependent repositioning to be diminished in both mutants (Fig. 4a, b, and Supplementary Fig. 5). The specificity of this observation was corroborated by deletion of the SAGA member *ADA2*, which impaired activation-dependent recruitment, as did the deletion of *NUP1*, which is consistent with its role in Sac3 localization¹⁸ (Fig. 4a, b, and Supplementary Fig. 5). By contrast, deleting nucleoporins *NUP2* and *NUP53* or the SAGA histone acetyltransferase *GCN5* had no effect (Fig. 4a, b, Supplementary Fig. 5, and data not shown). In addition, Mlp1 or Nup60, components otherwise involved in mRNA retention²⁰, have no direct function in the activation-dependent confinement of *GAL* genes (Supplementary Figs 5 and 6).

Activation-dependent changes in *GAL* positioning are not linked to *GAL1* mRNA transcription levels, which remain unaffected in *ada2Δ* and *nup1Δ* cells, and show only a modest decrease in *sac3Δ* cells (Fig. 4c). Moreover, by employing RNA-FISH we can show that the transcriptional activity of the *GAL* locus continues in *sus1Δ* and *sac3Δ* mutants unabated by the clearly intranuclear positioning of the active *GAL* loci (Supplementary Fig. 7). It indicates that although transcriptional activation is essential for the dynamic link to form, it is not sufficient because mRNA transcription alone does not mediate perinuclear confinement. Considering the direct biochemical interaction of the Sac3–Thp1–Cdc31 complex with both the SAGA complex (through Sus1) and the NPC (through Nup1), our data strongly support the formation of a direct and dynamic link between the activated gene and the NPC.

Taken together, the data presented here indicate that dynamic motility of genes within the nuclear space constitutes a determinant of regulation. Transcriptional factors responsible for gene activation concomitantly 'gate' genes by constraining the 3D motility of activated loci to defined nuclear positions. The resulting link of mRNA transcription and NPC export sites described here is likely to help the preferential processing and export of transcripts. The existence of intranuclear *GAL1* transcription sites in wild-type cells (Fig. 2b) indicates that *GAL* genes are first activated and then 'gated'. However, the reverse may be true for other loci, where the perinuclear position could dictate expression.

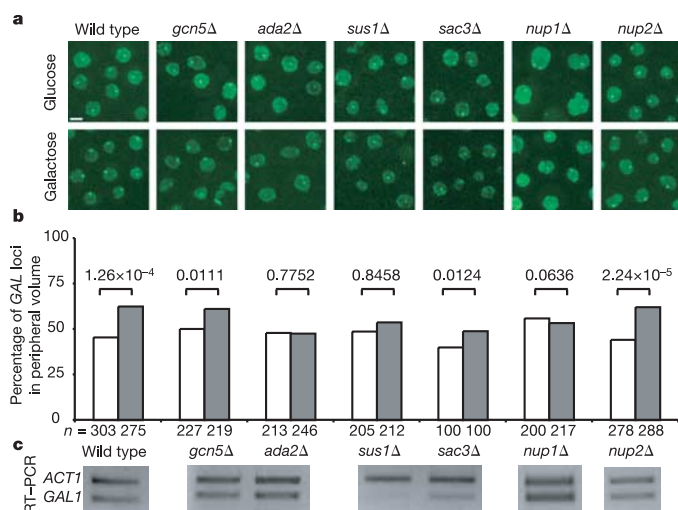


Figure 4 | Localization of *GAL* genes in SAGA and NPC mutants. **a**, *In vivo* 3D localizations of labelled *GAL* genes were performed for wild-type, *GCN5*-, *ADA2*-, *SUS1*-, *SAC3*-, *NUP1*- and *NUP2*-deleted cells (strains yGC105, yGC197, yGC196, yGC204, yGC205, yGC198 and yGC193, respectively). For each mutant, examples of 3D raw data are shown Z-projected in both glucose and galactose. Scale bar, 2 μ m. **b**, Bar graph indicating the proportion of *GAL* loci found in the peripheral volume (as defined in Supplementary Fig. 1). Pairwise comparisons of the distributions of R3Dp (see Methods) computed for each strain and growth condition were performed with Kolmogorov–Smirnov tests. The *P* value is indicated above the bracket for each test; *n* is the number of cells analysed. Open bars, glucose; filled bars, galactose. **c**, Level of *GAL1* transcripts in the galactose growth condition analysed in each mutant by RT–PCR followed by agarose-gel electrophoresis. *ACT1* transcripts are also included as a control.

METHODS

Microscopy. Live microscopy was performed with a Perkin-Elmer UltraView RS Nipkow-disk confocal system controlled with PE-viewer software (version 1.0.0.9) featuring a Zeiss 100 $\times$ objective (Plan Apo, 1.4 NA, oil immersion). The pixel size was 65.8 nm. For 3D static analysis, Z-stacks of 51 images with a 200-nm Z-step were used. The exposure time was 200 ms for GFP. For 3D dynamic analysis, 25-image Z-stacks with a Z-step of 250 nm and an acquisition time of 150 ms were taken every 4 s for 15 min (226 time points; 5,650 images).

RNA fluorescent *in situ* hybridization was performed as described²⁰.

Image analysis. Detection and tracking of GFP-tagged chromosomal loci and the NE in each Z-stack series were performed automatically with a dedicated software described in Supplementary Methods. After processing, the SPB labelled with Spc29–GFP was used as an internal reference to correct the Z-detection artefact (Supplementary Fig. 1a, b). Finally, the position in the nuclear volume of the detected locus is given by the relative 3D position (R3Dp), which is the ratio of (corrected distance between the centroid of the nucleus and the labelled locus) to (corrected radius). The nuclear space was divided into two concentric zones of equal volume (Supplementary Fig. 1c) and in each experiment the percentages of loci distributed between these two volumes computed over the cell sample of R3Dp are given. The cumulative distribution function (CDF), 3D graph and radial projection were plotted with Matlab software, version 7.0.4. Kolmogorov–Smirnov tests were performed and significance was determined with 95% confidence limits. Time-lapse experiments were analysed by computing the radMSD (Supplementary Fig. 3a) and the MSD (Supplementary Fig. 3b) with Matlab software, version 7.0.4.

Plasmids, strains, growth conditions, RT–PCR protocol and details of

quantitative image analysis and brownian motion analysis are described in Supplementary Methods.

Received 27 January; accepted 24 March 2006.

1. Spector, D. L. The dynamics of chromosome organization and gene regulation. *Annu. Rev. Biochem.* **72**, 573–608 (2003).
2. Andrusis, E. D., Neiman, A. M., Zappulla, D. C. & Sternglanz, R. Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* **394**, 592–595 (1998).
3. Casolari, J. M., Brown, C. R., Drubin, D. A., Rando, O. J. & Silver, P. A. Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. *Genes Dev.* **19**, 1188–1198 (2005).
4. Casolari, J. M. *et al.* Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* **117**, 427–439 (2004).
5. Brickner, J. H. & Walter, P. Gene recruitment of the activated INO1 locus to the nuclear membrane. *PLoS Biol.* **2**, e342 (2004).
6. Menon, B. B. *et al.* Reverse recruitment: the Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. *Proc. Natl Acad. Sci. USA* **102**, 5749–5754 (2005).
7. Blobel, G. Gene gating: a hypothesis. *Proc. Natl Acad. Sci. USA* **82**, 8527–8529 (1985).
8. Michaelis, C., Ciosk, R. & Nasmyth, K. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91**, 35–45 (1997).
9. Gotta, M. *et al.* The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type *Saccharomyces cerevisiae*. *J. Cell Biol.* **134**, 1349–1363 (1996).
10. Kimura, H., Sugaya, K. & Cook, P. R. The transcription cycle of RNA polymerase II in living cells. *J. Cell Biol.* **159**, 777–782 (2002).
11. Vazquez, J., Belmont, A. S. & Sedat, J. W. Multiple regimes of constrained chromosome motion are regulated in the interphase *Drosophila* nucleus. *Curr. Biol.* **11**, 1227–1239 (2001).
12. Havlin, S. & Ben-Avraham, D. Diffusion in disordered media. *Adv. Phys.* **51**, 187–292 (2002).
13. Bouchaud, J. P. & Georges, A. Anomalous diffusion in disordered media: statistical mechanics, models and physical applications. *Phys. Rep.* **195**, 127–293 (1990).
14. Huisinga, K. L. & Pugh, B. F. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol. Cell* **13**, 573–585 (2004).
15. Bhaumik, S. R. & Green, M. R. SAGA is an essential *in vivo* target of the yeast acidic activator Gal4p. *Genes Dev.* **15**, 1935–1945 (2001).
16. Larschan, E. & Winston, F. The *S. cerevisiae* SAGA complex functions *in vivo* as a coactivator for transcriptional activation by Gal4. *Genes Dev.* **15**, 1946–1956 (2001).
17. Fischer, T. *et al.* Yeast centrin Cdc31 is linked to the nuclear mRNA export machinery. *Nature Cell Biol.* **6**, 840–848 (2004).
18. Fischer, T. *et al.* The mRNA export machinery requires the novel Sac3p–Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. *EMBO J.* **21**, 5843–5852 (2002).
19. Rodríguez-Navarro, S. *et al.* Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* **116**, 75–86 (2004).
20. Galy, V. *et al.* Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. *Cell* **116**, 63–73 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Fagegaltier, E. Fabre, P. Therizol, B. Zhang, T. Fischer and all the members of the Nehrbass group for critical discussions, and A. Taddei and S. Gasser for sharing unpublished results. We thank the 'Plateforme d'imagerie dynamique' of the Pasteur Institute for providing access to the microscopy facilities. This work was supported by an ACI-BCMS grant. G.G.C. and A.G. were recipients of fellowships of the Ministère Français délégué à la Recherche et aux Nouvelles Technologies.

Author Contributions E.C.H. and U.N. conceived the project. G.G.C. did the experiments. S.R.-N. did the RT-PCR experiment. G.G.C. and O.G. conceived the image analysis protocols. A.G. and J.-C.O.-M. conceived the image analysis algorithms in collaboration with G.G.C. and O.G. A.G. implemented the image analysis algorithms. G.G.C. did the image processing and analysis. G.G.C. and C.Z. computed the MSD. G.G.C., O.G., C.Z., H.B. and A.L. interpreted the image analysis results. G.G.C., S.R.-N. and F.F.-F. provided the yeast strains and plasmid constructs. U.N. wrote the paper. F.F.-F. supervised G.G.C. J.-C.O.-M. supervised A.G. J.-C.O.-M., E.C.H. and U.N. are head of the laboratories participating in this work.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to U.N. (nehrbass@pasteur.fr) or E.C.H. (cg5@ix.urz.uni-heidelberg.de). Any requests for imaging software should be addressed to J.-C.O.-M. (jcolivo@pasteur.fr).

LETTERS

Nuclear pore association confers optimal expression levels for an inducible yeast gene

Angela Taddei^{1,2}, Griet Van Houwe³, Florence Hediger¹, Veronique Kalck¹, Fabien Cubizolles¹, Heiko Schober^{1,3} & Susan M. Gasser^{1,3}

The organization of the nucleus into subcompartments creates microenvironments that are thought to facilitate distinct nuclear functions¹. In budding yeast, regions of silent chromatin, such as those at telomeres and mating-type loci, cluster at the nuclear envelope creating zones that favour gene repression^{1,2}. Other reports indicate that gene transcription occurs at the nuclear periphery, apparently owing to association of the gene with nuclear pore complexes^{3–5}. Here we report that transcriptional activation of a subtelomeric gene, *HXX1* (hexokinase isoenzyme 1), by growth on a non-glucose carbon source led to its relocalization to nuclear pores. This relocation required the 3′ untranslated region (UTR), which is essential for efficient messenger RNA processing and export, consistent with an accompanying report⁶. However, activation of *HXX1* by an alternative pathway based on the transactivator VP16 moved the locus away from the nuclear periphery and abrogated the normal induction of *HXX1* by galactose. Notably, when we interfered with *HXX1* localization by either antagonizing or promoting association with the pore, transcript levels were reduced or enhanced, respectively. From this we conclude that nuclear position has an active role in determining optimal gene expression levels.

Several studies have reported that changes in the three-dimensional (3D) organization of the genome accompany changes in gene expression¹. In yeast, transcriptionally silent chromatin, including telomeres and silent mating-type loci, bind to the nuclear envelope through heterochromatin factors^{7–9}, whereas other genes associate with the nuclear periphery when activated^{3,4}. According to the ‘gene gating hypothesis’, this relocalization could facilitate mRNA export¹⁰. However, it is currently unknown whether the perinuclear positioning of genes is necessary for maximal expression levels.

We have examined this question by testing whether the activation of a budding yeast subtelomeric gene, *HXX1*, alters the location of the adjacent telomere (Tel6R), which is transcriptionally silent and peripherally positioned⁷. This locus was monitored in living cells by fluorescence microscopy using an array of *lac* operators¹¹ that were inserted upstream of the *HXX1* promoter, 13 kilobases (kb) from the terminal TG repeat. By expressing both a green fluorescent protein (GFP)–*LacI* repressor fusion and a GFP-tagged nucleoporin (GFP–Nup49)¹², we can monitor the position of the tagged locus by scoring its position relative to the nuclear periphery in a single focal plane of a 3D-image stack. A random position in a population of nuclei scores as 33% in each of three zones of equal surface. Values above or below 33% show enrichment or depletion, respectively (Fig. 1a)⁷.

Under conditions that repressed the *HXX1* promoter (that is, growth on glucose medium), this subtelomeric locus was adjacent to the nuclear envelope in 60% of the cells (Fig. 1b). When the promoter was activated by removing glucose (that is, growth on galactose

medium), *HXX1* transcript levels increased ~17-fold (Fig. 1e), as monitored by quantitative real-time polymerase chain reaction (PCR). As expression increased, the percentage of cells in which *HXX1* was found at the outermost edge of the nucleoplasm increased to >85% (Fig. 1b). In contrast, an internal gene whose transcription is insensitive to galactose (*PES4*; ref. 8), showed no change in position.

To rule out non-specific, galactose-induced shift of telomeric chromatin to the nuclear periphery, we induced *HXX1* in the presence of glucose by deleting its repressor, *HXX2* (refs 13, 14; Fig. 1c). Transcript amounts reached levels slightly higher than those monitored in wild-type cells on galactose, and again *HXX1* became more tightly associated with the nuclear periphery (Fig. 1d, e). This relocalization correlated strictly with *HXX1* activation, as it did not occur if the *HXX1* promoter and coding region were deleted (Fig. 1d).

Nuclear export factors have been proposed to link activated genes with the nuclear envelope⁵. Indeed, Sac3 and Sus1, both components of the transcription/export factor (TREX), are necessary for positioning the galactose-induced *GAL1* gene at the nuclear periphery⁶. Here we found that the deletion of the 3′ UTR of *HXX1* abolished its transcription-induced anchoring and reduced expression levels (Fig. 1d, e), suggesting that 3′-UTR-dependent mRNA processing influences the relocalization of the activated *HXX1* gene.

We next asked whether this transcription-coupled anchoring was related to other pathways of telomere anchoring. Previous studies showed that, under glucose growth conditions, the perinuclear positioning of native Tel6R was fully dependent on the conserved DNA-end-binding heterodimer, yKu70/yKu80 (ref. 7; Fig. 2a). However, in *yku70Δ* cells, the *HXX1* gene still shifted from a random to a peripheral position upon galactose induction (Fig. 2a; $P = 8 \times 10^{-7}$), arguing for a yKu-independent mechanism for tethering active genes.

Cross-linking studies³ suggest that transcription-coupled anchoring reflects association with nuclear pore complexes (NPCs). To test this, we used a *nup133*-deletion strain (*nup133Δ*) in which cyan fluorescent protein (CFP)-tagged nuclear pores cluster at one side of the nuclear rim (Fig. 2b). We scored the frequency of a complete superposition of GFP-tagged *HXX1* with the CFP-pore signal under repressed and inducing conditions: on glucose, the *HXX1* locus coincided very rarely with nuclear pores (4%), whereas activation of the gene increased colocalization of the fluorescent signals fourfold (to 17%; Fig. 2c).

This relatively low value for complete colocalization is likely to stem from the constant local motion of chromatin in living cells^{7,12}. Indeed, live 3D-tracking of the *HXX1* locus by confocal microscopy revealed a continuous diffusive movement of Tel6R within a radius of constraint of 0.6 μm in cells grown on glucose (Supplementary Fig. 2). When *HXX1* was induced, the locus remained dynamic,

¹Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland. ²UMR 218, Centre National de la Recherche Scientifique/Institut Curie-Section de Recherche, 26 rue d’Ulm, 75231 Paris Cedex 05, France. ³University of Geneva, NCCR Frontiers in Genetics, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland.

but moved in a more restricted peripheral zone, sliding laterally along the nuclear envelope (Fig. 2d). This mobility may reflect competition between the transcription-coupled anchoring of *HXK1* at nuclear pores and telomere-mediated anchoring at other sites (Fig. 2e).

To check whether the association of the activated *HXK1* gene with the nuclear envelope is an obligatory feature of active transcription,

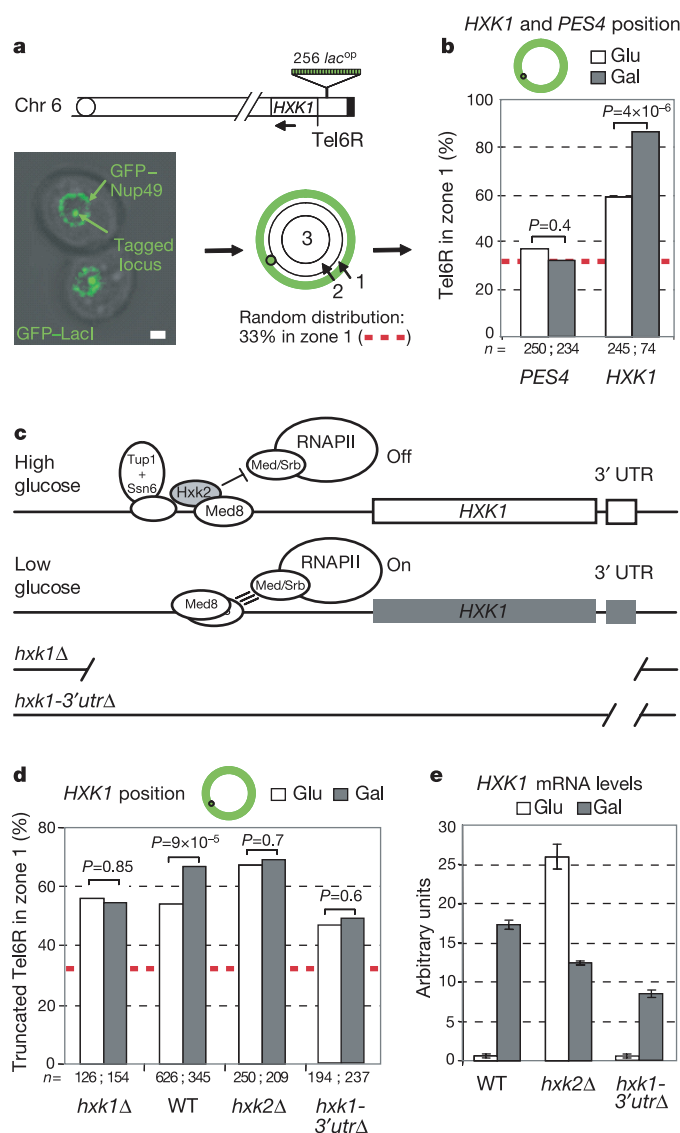


Figure 1 | Relocation of *HXK1* to the nuclear periphery on galactose requires the 3' UTR. **a**, Yeast cells expressing GFP-LacI and GFP-Nup49 fusions were tagged with 256 *lacOP* repeats upstream of *HXK1* and 13 kb from the TG repeats of Tel6R. *HXK1* subnuclear position was scored with respect to zone 1, representing 33% of the cross-sectional area of one focal plane of the nucleus. Scale bar, 1 μ m. Chr, chromosome. **b**, The percentage of spots are plotted for zone 1 (n = number of interphase cells) for strains bearing *lacOP* repeats near *HXK1* or *PES4* (ref. 8), in glucose (Glu) or galactose (Gal) media. *P*-values refer to a proportional test comparing indicated distributions. **c**, Schematic of *HXK1* gene regulation¹³ and the deletions relevant for panels **d** and **e**. **d**, As in **b**, for wild-type (WT) strain GA-1917 and isogenic derivatives *hxk1Δ* (GA-3328), *hxk2Δ* (GA-3329) or *hxk1-3'utrΔ* (GA-3520). In these strains, Tel6R is truncated, eliminating its subtelomeric X element⁷ (see Fig. 3a). **e**, *HXK1* mRNA levels were determined by quantitative real-time PCR with normalization by geomean to *NUP1* and *NUP159* transcripts²⁷. Error bars correspond to s.e.m. For the *hxk1-3'utrΔ* strain, GA-3520, grown either on glucose or galactose, internal sequence was used for the 3' primer rather than polyA, which gives a transcript level of 12.8 for the WT control on galactose.

we induced *HXK1* through an alternative pathway: the activation domain of the viral transcriptional activator VP16 (ref. 15) was fused to bacterial LexA and targeted to four LexA recognition sites inserted upstream of the *HXK1* promoter (Fig. 3a). A second gene, *ADE2*, was inserted immediately adjacent to the Tel6R TG₁₋₃ repeats to monitor the effects of LexA-VP16 on other subtelomeric genes. *HXK1* expression levels and *HXK1* position relative to NPCs were scored. As expected, targeting VP16 upstream of the *HXK1* promoter in glucose medium increased mRNA levels, this time 4.4-fold (Fig. 3b). We checked for the presence of RNA polymerase II in its elongation form by chromatin immunoprecipitation, and its level increased even more significantly than on galactose (Supplementary Fig. 1). Notably, this VP16-mediated activation was accompanied by a loss of perinuclear anchoring (Fig. 3c), although expressing LexA-VP16 did not affect telomeres that do not bear inserted LexA sites (see Tel8L,

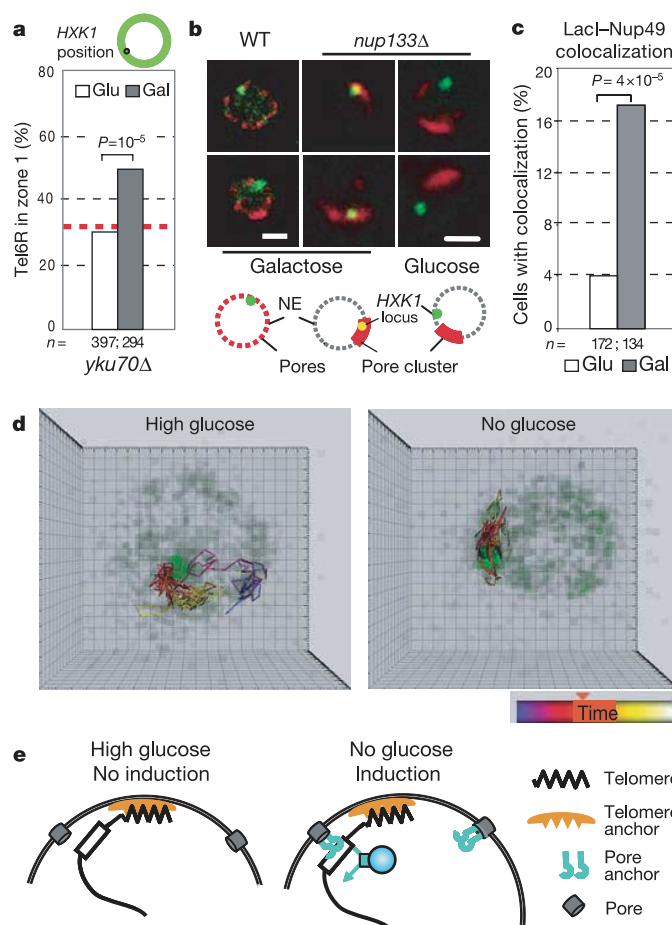


Figure 2 | *HXK1* colocalizes with nuclear pores in glucose-free medium. **a**, Bar graphs present the percentage of *HXK1* spots in zone 1 (see Fig. 1b) for strain GA-1489 (*yku70Δ*). **b**, Deconvolved equatorial confocal sections of nuclei from glucose- or galactose-grown wild-type or *nup133Δ* cells in which GFP-tagged *HXK1* is green and NPCs are red (CFP-Nup49). Scale bars, 1 μ m. NE, nuclear envelope. **c**, Quantification of complete signal overlap (colocalization) of *HXK1* with Nup49 (see **b**) (n =number of cells counted). **d**, Four-dimensional time-lapse imaging of wild-type cells bearing *lacOP*-tagged *HXK1* as described²⁶ on glucose- or galactose-containing media. Image stacks were collected at 1.5-s intervals for 7.5 min and tracking was performed with the Imaris Time program. The darker grey background shows the peripheral pore signal. The path of the locus is shown with a temporal colour code (blue represents early and white represents late time points). **e**, Schematic recapitulating *HXK1* activation and localization to nuclear pores in glucose-free medium. Ku-mediated telomere anchoring tethers silent *HXK1* to the periphery in glucose media, until transcription is induced and NPC association increases.

Fig. 3c). Moreover, the TG-proximal *ADE2* gene remained significantly repressed (Fig. 3b) in a variegated manner (Supplementary Fig. 3), indicating that *HXX1* transcription did not disrupt more distal telomeric heterochromatin¹⁶.

The displacement of Tel6R from the nuclear periphery may reflect VP16-recruited nucleosome remodellers or other transactivating factors, which provide a force sufficient to override the anchoring activity of telomeric chromatin^{7,8,17}. Monitoring Tel6R dynamics showed that VP16 binding enhanced chromatin movement and enlarged the radius of constraint within which the locus moved. This could displace Tel6R from the nuclear envelope (Supplementary Fig. 4).

The occurrence of transcription without perinuclear enrichment was not unique to activation by VP16. Similar results were obtained by inducing the subtelomeric *ADE2* gene by growth on adenine-deficient media. This internalization was statistically significant if anchoring was first weakened by eliminating one of the telomeric

anchors^{7,17} (*yku70Δ*; Fig. 3d). Thus, stable NPC association is not an absolute requirement for transcription in yeast, although our results do not rule out transient contacts between genes and pores during activation^{18,19}.

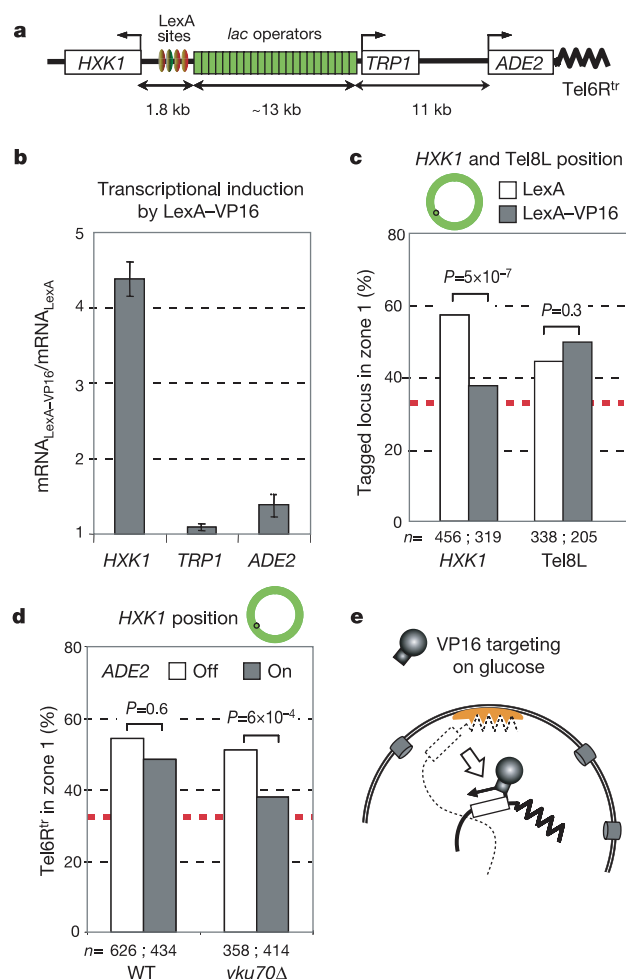


Figure 3 | Alternative modes of transcriptional activation displace Tel6R from the nuclear periphery. **a**, Organization of relevant subtelomeric elements in yeast strain GA-1917, in which Tel6R is truncated (Tel6R^{tr}) by the integration of TG-*ADE2* 5 kb from the native chromosomal end. Unlike native Tel6R, Tel6R^{tr} remains anchored by Sir4 and Esc1 in the absence of *yku70*¹⁷. **b**, Ratios of transcript levels in glucose-grown cells bearing LexA-VP16 or LexA alone determined by real-time PCR for *HXX1*, *TRP1* and *ADE2*. Error bars correspond to s.e.m. **c**, Subnuclear localization of *HXX1*/Tel6R^{tr} or Tel8L (no LexA sites) as in Fig. 1. Strains are as in **b**. **d**, Subnuclear position of *HXX1*/Tel6R^{tr} was determined as in Fig. 1 in GA-1917 (WT) or GA-1918 (*yku70Δ*) strains grown on synthetic medium with adenine ("Off") or without ("On"). **e**, Schematic recapitulating *HXX1* activation and internal localization on VP16 targeting to its promoter.

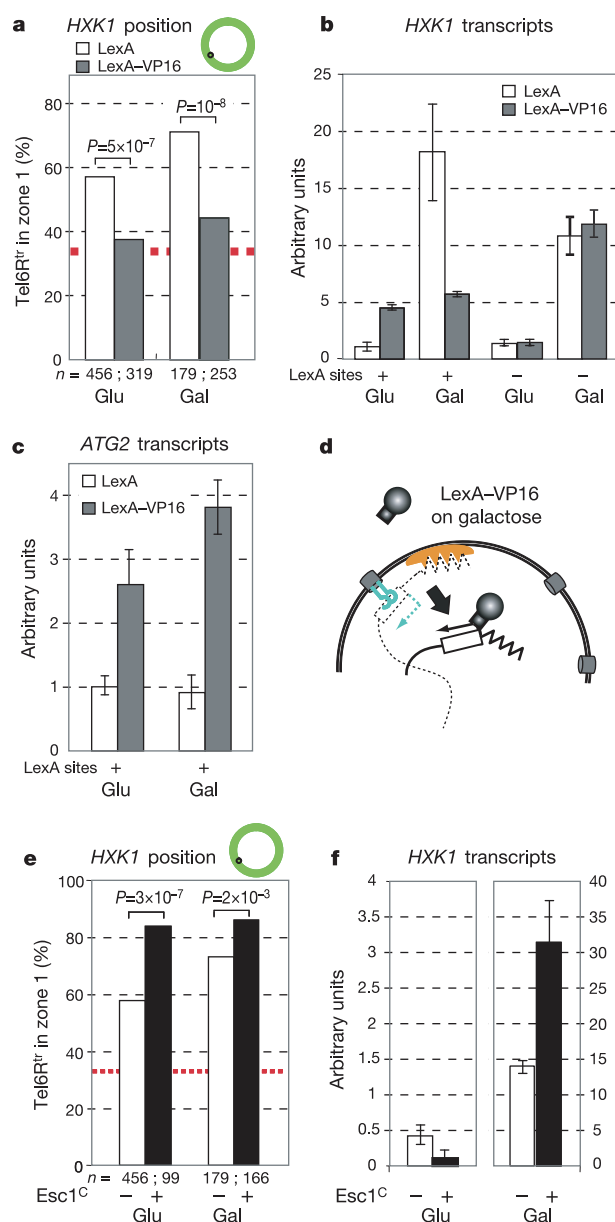


Figure 4 | Impairing or improving perinuclear anchoring modulates *HXX1* transcript levels on galactose. **a**, Subnuclear localization of *HXX1* was determined as in Fig. 1 in GA-1917 cells with or without VP16 on glucose or galactose. **b**, *HXX1* mRNA levels were determined as in Fig. 1e for isogenic strains bearing LexA sites near *HXX1* (GA-1917) or not (GA-1320). Error bars correspond to s.e.m. **c**, *ATG2* mRNA levels in GA-2070 cells bearing LexA binding sites and *lac*^{op} repeats downstream of *ATG2*, with or without VP16 on galactose or glucose. Error bars correspond to s.e.m. **d**, Schematic recapitulating *HXX1* activation and localization by two antagonistic modes of activation. Activation by VP16 on galactose shifts *HXX1* away from the nuclear periphery, and induction levels are reduced over levels on galactose alone. Similarly, a reduction in *HXX1* transcription and internalization was achieved on galactose when the subtelomeric *ADE2* gene was induced in a *yku70Δ* mutant (without adenine; data not shown). **e**, Subnuclear localization of *HXX1* was determined as in Fig. 1 in GA-1917 cells expressing LexA or LexA-Esc1^C on glucose or galactose. **f**, *HXX1* mRNA levels in cells expressing LexA or LexA-Esc1^C on glucose or galactose determined as in Fig. 1e. Right and left panels have different scales. Error bars correspond to s.e.m.

Our data thus far lead to two important conclusions. First, transcriptional activation can occur in yeast without a pronounced association with the nuclear envelope. Second, the position of an active gene depends, at least in part, on its mode of activation. We decided to test whether the observed pore association might be necessary for the high mRNA levels detected on galactose-containing medium by analysing transcript levels and gene position in the presence of both VP16 and galactose.

When *HXX1* was induced on galactose-containing medium in the presence of LexA–VP16, the tagged *HXX1* locus shifted inwards assuming a near-random distribution (Fig. 4a). Thus, the VP16 effect was epistatic to and dominant over the galactose-induced pore association (see Figs 1 and 2). Remarkably, under these conditions *HXX1* mRNA levels did not increase to the high level expected on galactose (~17-fold increase), but remained at the level conferred by LexA–VP16 on glucose (~4-fold above the control; Fig. 4b). Notably, the presence of LexA–VP16 had no effect on the galactose-induced level of *HXX1* transcription if the LexA binding sites were absent (Fig. 4b). Other galactose-induced genes, such as *GAL1* and *GLK1* (refs 13, 14), showed near-normal levels of induction in the same cells (Supplementary Table 3), ruling out indirect effects of LexA–VP16 on galactose-regulated gene activity (that is, 'silencing'²⁰). Finally, the galactose effect was specific for LexA–VP16-activated *HXX1*, because galactose did not impair the activation of another locus to which the fusion was targeted (*ATG2*; Fig. 4c). Thus, the binding of LexA–VP16 upstream of the *HXX1* promoter specifically abrogated the normal level of *HXX1* expression on galactose (Fig. 4b). Because this reduced expression level was accompanied by a marked dissociation of *HXX1* from the nuclear periphery (Fig. 4d), we propose that the association of *HXX1* with nuclear pores is required for maximal expression.

As a corollary to this observation, we next tested whether improved anchoring could enhance *HXX1* expression even further. To achieve this, we targeted the anchoring domain of the nuclear-envelope-associated protein Esc1 (LexA–Esc1^C)⁸ to the LexA sites upstream of *HXX1*, improving significantly the association of *HXX1* with the nuclear periphery in both glucose and galactose media (Fig. 4e). This enhanced perinuclear anchoring led to a hyper-repression of *HXX1* on glucose; however, on galactose, the induced *HXX1* mRNA levels reached values twice as high as those monitored on galactose in the absence of the Esc1 anchor (Fig. 4f). This argues that increased association with sites near pores enhances *HXX1* activation on galactose, presumably by facilitating access to NPC sites, just as antagonizing this interaction reduced expression levels.

We conclude that although transcription-associated NPC anchoring is not necessary for all modes of transcriptional activation in yeast, for *HXX1* such association maximizes expression levels. We note that Med8 (ref. 13), which is thought to bind the *HXX1* promoter, interacts with the integral pore protein Nup116 (ref. 21), and can be recovered in a complex with the nuclear-pore-associated factor Mlp1 (ref. 22). These interactions may help anchor the activated *HXX1* promoter to the NPC. However, mRNA processing/export factors also contribute to this transcription-induced pore-anchoring, as the relocalization of *GAL1* to NPCs required Sac3 and Sus1 (refs 5, 6), and deletion of the *HXX1* 3' UTR abrogated the enrichment of active *HXX1* at the pore (Fig. 1). Consistent with these findings, the elongation-specific form of RNA polymerase II was as abundant on *HXX1* in the presence of VP16 as it was on galactose (Supplementary Fig. 1), even though the accumulated level of *HXX1* mRNA is fourfold lower. This argues that, in this instance, RNA polymerase II loading is not the limiting factor that determines final transcript levels, and supports the proposal that NPC association maximizes expression levels by facilitating mRNA export or processing^{5,6}. This may reflect cooperation between promoters and 3' sequences, which were proposed to form active gene loops in yeast²³.

In conclusion, different modes of gene activation can lead to different subnuclear positions. Because interfering with gene relocation

on galactose modified *HXX1* expression level, we propose that sub-nuclear position has an active role in determining optimal gene expression levels.

Note added in proof: Consistent with our results, another study²⁸ has observed galactose-dependent accessibility of the *HXX1* promoter to a Nup2 fusion protein.

METHODS

Microscopy. For *in vivo* position analysis, a Metamorph-driven Olympus IX70 microscope was used to capture 21 image stacks of 0.2-μm step size. The zonal position of the GFP spot was determined on one focal section⁷. Deconvolution was performed using Metamorph software⁸.

Time-lapse imaging was performed using a Zeiss LSM510 confocal microscope as described²⁴ on strains GA-1459 (ref. 12) (pixel size: 100 nm×100 nm; imaging intervals: 1.5 s; z-settings: 6 optical slices at 450 nm spacing)^{25,26}. Each cell was monitored for 7.5 min, and subsequently scored for cell division timing.

Tracking and movie analysis. Four-dimensional (4D) tracking was performed with the Imaris Time program (Bitplane). The tagged locus was tracked using the SpotTracker plug-in²⁵ (ImageJ; <http://rsb.info.nih.gov/ij/index.html>). MSD (mean-square displacement) analysis is described in Supplementary Information.

Transcript analysis. Transcript analysis and quantification were performed on at least four colonies for each condition by reverse transcription and quantitative real-time PCR. Primers for *ADE2* and *TRP1* mRNA analysis were designed to detect the wild-type, but not the mutated, alleles (that is, *ade2-1* and *trp1-1*) present at endogenous loci (Supplementary Table 2). Values were normalized by geomean to *NUP1* and *NUP159* for each condition²⁷. The error is the standard error of the mean of all measurements.

Received 30 January; accepted 28 April 2006.

1. Taddei, A., Hediger, F., Neumann, F. R. & Gasser, S. M. The function of nuclear architecture: a genetic approach. *Annu. Rev. Genet.* **38**, 305–345 (2004).
2. Andralis, E. D., Neiman, A. M., Zappulla, D. C. & Sternglanz, R. Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* **395**, 592–595 (1998).
3. Casolari, J. M. *et al.* Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* **117**, 427–439 (2004).
4. Brickner, J. H. & Walter, P. Gene recruitment of the activated *INO1* locus to the nuclear membrane. *PLoS Biol.* **2**, e342 (2004).
5. Rodríguez-Navarro, S. *et al.* Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* **116**, 75–86 (2004).
6. Cabal, G. G. *et al.* SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature* doi:10.1038/nature04752 (this issue).
7. Hediger, F., Neumann, F. R., Van Houwe, G., Dubrana, K. & Gasser, S. M. Live imaging of telomeres. yKu and Sir proteins define redundant telomere-anchoring pathways in yeast. *Curr. Biol.* **12**, 2076–2089 (2002).
8. Taddei, A., Hediger, F., Neumann, F. R., Bauer, C. & Gasser, S. M. Separation of silencing from perinuclear anchoring functions in yeast Ku80, Sir4 and Esc1 proteins. *EMBO J.* **23**, 1301–1312 (2004).
9. Gotta, M. *et al.* The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type *Saccharomyces cerevisiae*. *J. Cell Biol.* **134**, 1349–1363 (1996).
10. Blobel, G. Gene gating: a hypothesis. *Proc. Natl Acad. Sci. USA* **82**, 8527–8529 (1985).
11. Robinett, C. *et al.* *In vivo* localization of DNA sequences and visualization of large scale chromatin organization using lac-operator/repressor recognition. *J. Cell Biol.* **135**, 1685–1700 (1996).
12. Heun, P., Laroche, T., Shimada, K., Furrer, P. & Gasser, S. M. Chromosome dynamics in the yeast interphase nucleus. *Science* **294**, 2181–2186 (2001).
13. de la Cera, T., Herrero, P., Moreno-Herrero, F., Chaves, R. S. & Moreno, F. Mediator factor Med8p interacts with the hexokinase 2: implication in the glucose signalling pathway of *Saccharomyces cerevisiae*. *J. Mol. Biol.* **319**, 703–714 (2002).
14. Rodríguez, A., de la Cera, T., Herrero, P. & Moreno, F. The hexokinase 2 protein regulates the expression of the *GLK1*, *HXX1* and *HXX2* genes of *Saccharomyces cerevisiae*. *Biochem. J.* **355**, 625–631 (2001).
15. Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. GAL4–VP16 is an unusually potent transcriptional activator. *Nature* **335**, 563–564 (1988).
16. Tham, W. H., Wyithe, J. S., Ko Ferrigno, P., Silver, P. A. & Zakian, V. A. Localization of yeast telomeres to the nuclear periphery is separable from transcriptional repression and telomere stability functions. *Mol. Cell* **8**, 189–199 (2001).
17. Hediger, F., Berthiau, A. S., van Houwe, G., Gilson, E. & Gasser, S. M. Subtelomeric factors antagonize telomere anchoring and Tel1-independent telomere length regulation. *EMBO J.* **25**, 857–867 (2006).

18. Ishii, K., Arib, G., Lin, C., Van Houwe, G. & Laemmli, U. K. Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell* **109**, 551–562 (2002).
19. Menon, B. B. *et al.* Reverse recruitment: the Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. *Proc. Natl Acad. Sci. USA* **102**, 5749–5754 (2005).
20. Gill, G. & Ptashne, M. Negative effect of the transcriptional activator GAL4. *Nature* **334**, 721–724 (1988).
21. Ito, T. *et al.* A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl Acad. Sci. USA* **98**, 4569–4574 (2001).
22. Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
23. O'Sullivan, J. M. *et al.* Gene loops juxtapose promoters and terminators in yeast. *Nature Genet.* **36**, 1014–1018 (2004).
24. Hediger, F., Taddei, A., Neumann, F. R. & Gasser, S. M. Methods for visualizing chromatin dynamics in living yeast. *Methods Enzymol.* **375**, 345–365 (2004).
25. Sage, D., Neumann, F. R., Hediger, F., Gasser, S. M. & Unser, M. Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics. *IEEE Trans. Image Process.* **14**, 1372–1383 (2005).
26. Gartenberg, M. R., Neumann, F. R., Laroche, T., Blaszczyk, M. & Gasser, S. M. Sir-mediated repression can occur independently of chromosomal and subnuclear contexts. *Cell* **119**, 955–967 (2004).
27. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, research0034.1–research0034.11 (2002).
28. Schmid, M. *et al.* Nup-PI: The nucleopore-promoter interaction of genes in yeast. *Mol Cell* **21**, 379–391 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Laroche, the imaging platform of the FMI, and the genomics platform of the NCCR for assistance, and E. Heard, A. Peters, G. Almouzni, D. Schübeler and M. Gartenberg for helpful suggestions, as well as G. Cabal and U. Nehrbass for sharing unpublished results. Our research is supported by the Swiss National Science Foundation, the NCCR programme 'Frontiers in Genetics', and the Novartis Research Foundation.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.G. (susan.gasser@fmi.ch).

naturejobs

**THE CAREERS
MAGAZINE FOR
SCIENTISTS**

I will always associate graduate school with pool. I shot as much as I could between research, writing and teaching — perhaps trying to get away from academic, career and personal concerns, or perhaps trying to let my subconscious work on the latest, seemingly intractable problem.

Sometimes my opponents — also seeking temporary exile from academic life — would imitate the cocky, narcissistic Tom Cruise character from *The Color of Money*. Inevitably, someone would sing the refrain from that movie's theme song, Eric Clapton's 'It's in the Way You Use it'. Especially when someone (usually me) missed an easy shot.

That song was a joke then, but it has relevance now, when smaller institutions try to compete with more established hubs such as Boston, San Francisco or Cambridge. The sentiment is the same for grad students and postdocs who bemoan the lack of job opportunities outside the lab, or a dearth of general career information. Scientists in Boston are lucky enough not to have this problem; the number of universities, research hospitals and biotech firms close by make options and collaborations readily available (see page 780). But you don't need to live and work in Boston to form an industrial collaboration, attend symposia at a nearby institution or take courses to pick up new skills. There are often local opportunities, and if connections don't exist between institutions or sectors, you could always start an organization to create them — even over long distances.

In *The Color of Money*, not taking advantage of his resources eventually left the Tom Cruise character snookered by his hustler mentor, Paul Newman. Scientists who don't tap into resources, be they local or otherwise, may find themselves similarly behind the eight ball.

Paul Smaglik, *Naturejobs* editor

CONTACTS

Publisher: Ben Crowe

Editor: Paul Smaglik

Assistant Editor: Gene Russo

US Head Office, New York

75 Varick Street, 9th Floor,

New York, NY 10013-1917

Tel: +1 800 989 7718

Fax: +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager/Corporations:

Peter Bless

Classified Sales Representatives

Tel: +1 800 989 7718

East USA/Canada: Andrew Bennie

NIH/Maryland/New York/

Pennsylvania: Shelley Cohen

San Francisco Office

Classified Sales Representative:

Michaela Bjorkman

West USA/West Corp. Canada

225 Bush Street, Suite 1453

San Francisco, CA 94104

Tel: +1 415 781 3803

Fax: +1 415 781 3805

e-mail: m.bjorkman@naturesf.com

European Head Office, London

The Macmillan Building,

4 Crinan Street, London N1 9XW, UK

Tel: +44 (0) 20 7843 4961

Fax: +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

European Sales Manager:

Andy Douglas (4975)

Advertising Production Manager:

Stephen Russell

To send materials use London

address above.

Tel: +44 (0) 20 7843 4816

Fax: +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Catherine Alexander

European Satellite Office

Patrick Phelan

e-mail: p.phelan@nature.com

Japan Head Office, Tokyo

Chiyoda Building,

2-37 Ichigayatamachi,

Shinjuku-ku,

Tokyo 162-0843

Tel: +81 3 3267 8751

Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:

Ayako Watanabe

e-mail: a.watanabe@natureasia.com

Bound for Boston

With its top academic institutions, Boston has long been a Mecca for biotechnology companies. But the demand for more academic input into industrial science is luring top drug companies there too, says **Corie Lok.**

A three-year-old, ten-storey modern glass building housing two of Harvard Medical School's basic research departments and several clinical programmes sits on the school's Boston campus. Its near-twin, a two-year-old, eleven-storey glass building housing Merck Research Laboratories stands next door.

Merck didn't set out to copy Harvard's neighbouring building, but it is fitting that the two buildings look similar. Merck's Boston laboratories are trying not just to be closer to Harvard, but also to be more like Harvard. That means recruiting scientists from Harvard and other nearby institutions and collaborating more with local academic and biotech researchers. "Boston-Cambridge is ideal because of the density of outstanding institutions," says Lex Van der Ploeg, the head of Merck's Boston labs.

Merck is one of several pharmaceutical companies that have realized Boston is the place to be. Indeed, more funding from the National Institutes of Health goes to Boston and Cambridge-based institutions (more than \$1.5 billion in 2004) than to those in any other US city. And Boston-Cambridge can boast perhaps the largest concentration of biotech companies of any US urban centre.

The drug companies hope that being close to so many potential academic and biotech collaborators will help them reinvigorate drug development. "It's no longer sufficient to be based in New Jersey to negotiate deals," says Kenneth Kaitin, director of the Tufts Center for the Study of Drug Development, also in Boston. The pursuit of new ideas and technologies from academia and biotech has gotten so competitive that drug companies "literally have to be in the same city to get early access to those assets."

Pharmaceutical companies have not been shy about moving next door to the universities. In 2002, Novartis moved its research headquarters from Basel, Switzerland, to Cambridge. One of its two buildings is across the street from the Whitehead Institute for Biomedical Research and the McGovern Institute for Brain Research at the Massachusetts Institute of Technology (MIT). The Pfizer Research Technology Center, which opened in 1999, is located just off the westernmost tip of the MIT campus. Wyeth Research has a site three subway stops away from Harvard University. In Cambridge, academic and industry collaborators can easily take a 15-minute stroll (or less) to meet up and talk science.

The arrival of the pharmaceutical research centres has opened up new job opportunities for scientists in Boston. "It's been very positive having them here," says Una Ryan, chair of the Massachusetts Biotech Council. "There's a sense that you'll have something to rely on if the first ten jobs don't work out."



OWAKI-KULLA/CORBIS

Although there was some initial fear in the community that smaller companies and even academic institutions wouldn't be able to attract workers lured away by the drug industry's high salaries, that fear for the most part has waned, says Ryan. In fact, the presence of the big drug companies can help enrich the skills of scientists in Boston and could even draw more scientists to the Boston area, says David Altshuler, director of the medical and population genetics programme at the Broad Institute. "One of the best job-training programmes that we've had at the Broad Institute is the local biotech and pharmaceutical industry," says Altshuler. He cites examples of people who used to work at Massachusetts General Hospital (where he is also based) or the Whitehead Institute/MIT Center for Genome Research (a precursor to the Broad Institute), but then moved to industry. After they'd gained new technical or management skills, they were hired back by the Broad for those new skills. "We recruit from companies as much as companies come and try to recruit from us," says Altshuler.

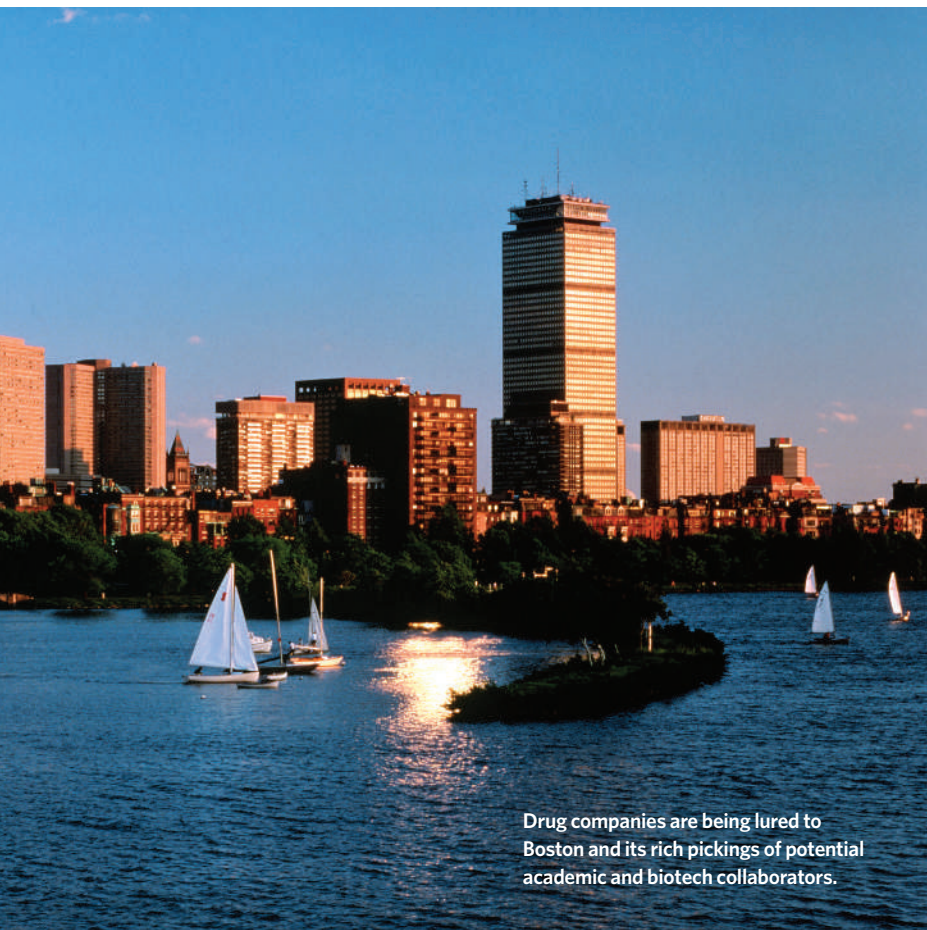
Hiring sprees

Still, competition for top scientists can be intense. Recruitment advertisements from companies such as Merck and Novartis are routinely found on the subway trains and stations that serve MIT and Harvard. Wyeth's Cambridge site reported an increase in turnover among its scientists last year, which was probably due to their competitors moving in and embarking on hiring sprees. Recruiting talent has been especially challenging for smaller, lesser-known firms such as Organon, a Dutch drug company that opened a

Merck is one of the many companies that has realized Boston is the place to be.

MERCK





Drug companies are being lured to Boston and its rich pickings of potential academic and biotech collaborators.

research centre last year a few blocks away from MIT. "There's a lot of competition but the pool is large," says Wiebe Olijve, who heads Organon's Cambridge centre.

Whereas most pharmaceutical companies established their presence in Boston–Cambridge by adding a Boston research centre to their list of sites around the world, Novartis went a step further by placing its research leadership in Cambridge. Like Merck, Novartis is trying to emulate its academic and biotech neighbours. To head the new Cambridge enterprise, it hired a Boston academic, Mark Fishman, who was the chief of cardiology and director of the Cardiovascular Research Center at the Massachusetts General Hospital before joining Novartis. Four of its 16 scientific leaders were recruited directly from Harvard or its affiliated research hospitals.

Since early 2003, Novartis's Cambridge labs have hired about 1,000 people; most are from the Boston area and many had fewer than five years of industry experience. To drive its local recruitment, the company goes to more than 20 local career fairs each year, even ones at smaller colleges, says Terry Gallagher, Novartis's head of staffing in Cambridge. "We want to be entrenched in the community," says Gallagher.

Merck's Boston labs opened in 2004. The labs employ about 200 scientists, and there are plans to hire 50 more this year and 100 more employees next year. Van der Ploeg says his research budget and recruitment plans weren't affected by the cutbacks at other Merck sites, which resulted in the layoffs of 7,000 workers worldwide last year. Nor has the research or recruitment at Boston been affected by the thousands of pending lawsuits surrounding the

withdrawal of its painkiller Vioxx, he says.

To help fuel its recruitment efforts, Merck has taken advantage of its proximity to academic centres. For example, Nancy Kohl, senior director of cancer biology and therapeutics, says she has had dozens of young, local academic scientists walk over to the Merck building to chat informally with her and others about what it's like to work there before deciding whether to apply for a position. Kohl got to know many of these scientists initially through research collaborations. Of the 13 recently hired PhDs in her department, four were from the Boston area, either from academia or from biotech companies.

Proximity helps

In addition to recruiting, pharmaceutical companies have come to Boston to establish ties with the local academic and biotech community. For example, Merck's Boston labs, unlike the company's other research sites, holds career fairs, symposia and seminars that are open to local academic scientists. They also host open-house days for students, and events for investors. The company has begun more than ten new collaborations with local academic groups and companies since opening the Boston labs.

Not long after moving to Cambridge, Novartis announced a collaboration with the Broad Institute to uncover the genetic roots of type 2 diabetes. The team is close to releasing its first set of results on the web, which will be freely available to scientists. Being so close together means team members from both organizations can meet weekly, says Altshuler, who is one of the principal investigators on the project. But the proximity doesn't just enable more frequent meetings, he says. "You end up with a different nature of relationship." There's a sense of a higher level of commitment to making the collaborations work, since it's in everyone's long-term interest to maintain good neighbourly relations, says Altshuler.

And these kinds of collaborations with industry can open doors to young academic scientists. Wyeth Research in Cambridge hired Robert Martinez in 2001 as a staff scientist after he had spent four years as a postdoc at the Dana-Farber Cancer Institute in Boston. Martinez started his postdoc intending to pursue a career in academia, largely because it was the only career option he really knew. But being in Boston, he got exposure to industry research. Although Martinez didn't collaborate directly with industry as part of his postdoctoral work, he was invited by his principal investigator, cancer researcher Todd Golub, to attend meetings with local companies that other members of the lab were working with. "It opened my eyes," says Martinez. Now at Wyeth, Martinez says he still attends seminars at academic institutions like the Whitehead Institute and the Dana-Farber Cancer Institute. He has even had bioinformatics experts from the Broad Institute come to train him and his colleagues on software the Broad has developed and made freely available.

Originally from the Caribbean, Martinez had once thought that he would work in California or Hawaii for the warm weather, but doesn't seem to regret his choice of climate. "I wanted to go to a place that would foster the advancement of my career," says Martinez. "Boston is certainly a place to do that."

Corie Lok is the editor of Nature Network Boston.
<http://beta.network.nature.com/boston>



Lex van der Ploeg and Nancy Kohl of Merck are reaping the benefits of their close proximity to academic centres in Boston.



MOVERS

Bruce Jones, chief scientist, United States Geological Survey, Washington DC



2003–2005 Senior scientist, US Environmental Protection Agency, Las Vegas, Nevada
1998–2003 Branch chief of landscape ecology, US Environmental Protection Agency, Las Vegas, Nevada
1992–1998 Technical coordinator for landscape ecology, US Environmental Protection Agency, Las Vegas, Nevada

When famed herpetologist and curator at the Bronx Zoo, Raymond Ditmars, entertained with stories of komodo dragons and their far-flung habitats, a seven-year-old Bruce Jones was hooked. "My life-long fascination with biology and biogeography began then," he says.

Snakes, however, charmed him into federal government. After getting his MSc in biology at New Mexico State University, Jones took a break from school to conduct habitat surveys of reptiles and amphibians for the Bureau of Land Management. From there, he listed endangered species for the US Fish and Wildlife Service.

He then moved to the US Environmental Protection Agency (EPA) to head an ecosystem-monitoring initiative called the Environmental Monitoring and Assessment Program. Eager to protect ecological systems as opposed to individual species, the project's success allowed him to establish a landscape-monitoring programme.

His pursuit of a PhD in environmental biology proved a turning point. "I got the brain boost that comes from debating theories without the repercussions of a federal title," he says. His biggest challenge remains working within the constraints of government science, whose policy focus sometimes falls short of academic-quality research.

Jones's work has helped change the EPA's focus from simply enforcing regulation on single eco-stressors to an integrated approach to environmental and human health. Using satellite imagery and geographic information science, his group published a landmark report on the ecological condition of the mid-Atlantic region of the United States.

One of his great successes involved a land-use change assessment — Jones's research showed that spending \$500 million to protect an upstate watershed, and therefore water quality, would save New York City the \$7 billion to build a filtration system. The oft-cited work is one of the first instances of placing economic value on services freely provided by ecosystems, such as water purification. He has since been involved in a NATO project with Europe and Australia to assess watershed conditions across the continents. "It's been a remarkable ride," he says.

Most recently, Jones left the EPA to become chief scientist for the US Geological Survey. He hopes to move the organization past mapping and into the science of spatial analysis and ecosystem vulnerability assessments. Jones says his first priority is to build a strong network of colleagues, his most important natural resource.

Virginia Gewin

BRICKS & MORTAR

Nurturing bioincubators in the north

After opening in November of last year, the new Sheffield Bioincubator has high hopes of contributing to bioscience proliferation in northern Britain. The facility follows several bioincubators that have already sprouted up in the region (*Nature* **425**, 430–433; 2003).

"It reflects strength of healthcare technology and the bioscience sector in the region," says Mark Tock, the Sheffield Bioincubator manager. "In the past, the northern regions hadn't been able to receive as much investment. It was largely a perception problem." Now, Tock suggests that "bioscience is booming under a steel sky".

The United Kingdom has a total of 17 bioincubators. Including Sheffield, seven are located in northern Britain — the York Bioincubator, the Institute of Pharmaceutical Innovation in Bradford, the Manchester Incubator, the MerseyBio Business Incubator in Liverpool, Biocity in Nottingham, and the Roslin Bioincubator. An eighth, the Leeds Bioincubator, will open for business in July 2007.

Twenty scientists and support staff currently occupy the building, says Tock. The incubator's 16 labs and 16 offices, though, have the potential to accommodate up to 150 staff. The incubator started with £6.9 million in investments from the University of

Sheffield and European Regional Development Funds, which were delivered through the South Yorkshire Objective 1 Programme.

The facility's first tenant, a company called Biofusion, specializes in the commercialization of intellectual property for life sciences, and, in particular, science generated at the University of Sheffield. One Biofusion spin-off diagnostic company, Lifestyle Choices, develops menopause and female fertility products. Another early tenant, ARC BioServ, provides custom production biomolecules, such as monoclonal and polyclonal antibodies. In July, the South Yorkshire Bioscience Enterprise Network will join the bioincubator.

The Sheffield facility also has a 'virtual network' to serve companies with less mature ideas that seek the benefit of the bioincubator's expertise. In return for a fee, companies can use the bioincubator as a professional front, taking advantage of a telephone answering service and meeting facilities. Virtual tenants can log on to the bioincubator's community website remotely. "It enables younger companies, and companies looking to expand into the UK and EU, to plug in to our community," Tok says.

Gene Russo

GRADUATE JOURNAL

Einstein's secret diary?

Routine pleasures get me through the day — that, and the grandiose dreams of a scientist in training. Welcome distractions pad out the path to a higher degree.

When I awake I'm slightly panic-stricken — have I slept in? A minute later my alarm goes off. I relax and go back to sleep. At work I check my e-mail every 20 minutes. Sometimes I become embroiled in esoteric arguments. I print off lots of articles and pile them on my desk with the hopeless intention of reading them all. I make sure to minimize the television schedule as my supervisor walks past. I scurry past her door a few times in my lab coat in the hope I look busy.

Often, I decide that I can't go on another minute without rating my favourite songs in iTunes or tidying my stationery drawer. I daydream that my lab mates look like the cast of *ER*. Occasionally I envisage myself making a major discovery and winning a Nobel prize.

When my supervisor heads home I check the latest celebrity gossip and turn the music up loud in defiance. Lab life rocks! When a sad song comes on I lapse into self-doubt: why aren't my experiments working? Am I the worst graduate student in the world? Early evening I go for a run. Afterwards I meet some fellow students and have a couple of drinks. I leave early 'to work'. Once home, I watch TV. In bed as I am about to nod off, I think of that great experiment I must do. I fall asleep dreaming of my Nobel prize acceptance speech.

Mhairi Dupré is a first-year PhD student in evolutionary developmental biology at the University of Oxford, UK.

The computifful game

A sporting chance.

Paul Steven Miller

Come join me as I bounce about the grassy soccer pitch of the Camp Nou, Barcelona. Climb inside my circuits. Climb inside my thoughts.

I wait impatiently in the centre circle, trapped beneath the right studded boot of the Pelé Mark VII. There are a million fans in the stadium, all anticipating the start of the match, roaring down from the tiers that stretch to the same height as the Eiffel tower.

As a child, my creator met the real Pelé, the greatest soccer player of the previous century. That was the beginning of his grand design.

I know this because my creator tells me everything. He has filled me with pinprick ears and eyes but given me no movable parts or mouth. "You're perfect," he often says, while tinkering in his workshop. "You listen and watch but you never tell. No, you never tell."

My creator's notoriety came not just from his genius but also from his prescience.

By the late 2040s, doping in sport was rife. Winning at any cost was the only maxim, and catalytic antibodies were the tonic. Based on variations of natural antibodies, they could be engineered to catalyse any aspect of biological function, and so enhance human performance. They were a doper's dream. Just as with the natural antibodies of the immune system, an almost infinite array of catalytic variants could be generated. So by the time the Antidoping Federation had developed an assay to detect one catalytic antibody the dopers had already designed another with the same function — but with a different structure to escape detection.

The Antidoping Federation had been rendered impotent. Yet it was the insurers who were the main problem. Many of the catalytic antibodies overworked the body, stretching the limits of biological endurance. Players were increasingly picking up career-ending injuries and demanding compensation. With no easy way of proving whether or not a player had been doped with catalytic antibodies at the time of injury, the insurers were more

reluctant than ever to pay up. Eventually, with no simple solution to the fiasco in sight, the insurers reneged on all their sports contracts.

No insurance meant no sport. It was illegal to participate in professional sports without a protective compensation policy: no one else could stand the litigation.

Ordinary citizens were in uproar, and so were the media. How were they going to fill the back pages of the newspapers? Chess? Lawn bowls?

But my creator was pleased. He had foreseen. He was prepared. He unveiled his robotic

have been immortalized in a new form. Female fans were also quietly impressed. Some of the robots in once male-dominated sports were modelled on the female physique: the soccerbot David Beckyham, the basketbot, Michelle Jordan, even the basebot, Babe Ruth. "This time around he really is a babe!" my master often jokes.

Besides, once the robosports got under way, most sports fans really couldn't tell the difference: robot technology was that good. And the algorithms for realistic decision-making in sports had existed in computer games for decades. The robots could even act like humans in the pre- and post-match interviews. David Beckyham was given a cockney accent and programmed to say 'at the end of the day' three times every sentence.

But the roboplayers aren't really smart. Not like me.

I am the culmination of my master's plans: an artificially intelligent football. He says it's because he has no real friends. He's a genius that no one understands. He's lonely. But I listen, always, and, mostly, I understand.

His lofty status and his perceived probity allow him to ensure I'm used in certain soccer matches. He knows I like nothing more than to be passed around the pitch.

Ah, yes! I want the match to start. An expectant silence is building in the stadium.

But things are different today. I'm excited but I also feel nervous. Last night in his workshop, my master tampered with one of the synthetic strips on my surface and inserted a micro-device underneath.

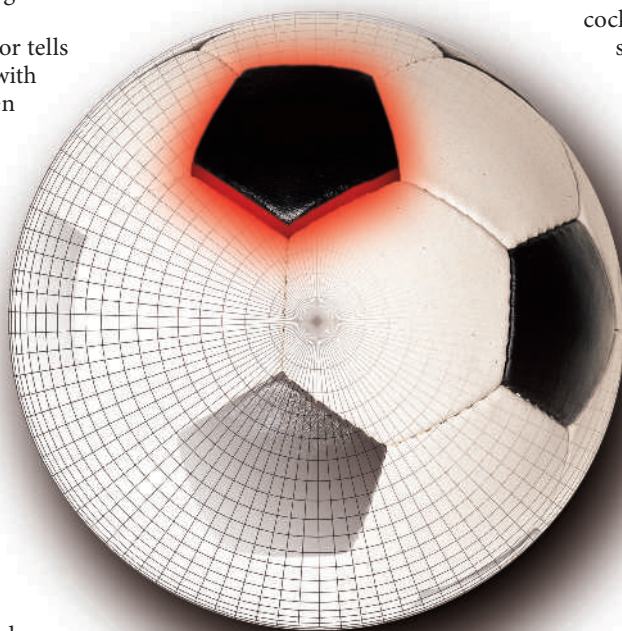
"With this," he said, leaning close to me, "I can remotely manipulate your flight-path during the game."

"Isn't that cheating?" he then said, speaking on my behalf.

"Why of course, yes!" he answered, laughing. "You won't tell anyone will you?"

No, I never tell. ■

Besides being an avid footy fan, Paul Steven Miller is a scientific researcher and obsessive sci-fi writer, seeking publication for his first novel *The Humanoid Foundation*.



doppelgangers to the world. He had been building a secret army of sportsbots, one to represent each great competitor in history.

People laughed at first. The media didn't: they seized the moment, hailing the robots as the solution, at least until the honour of human sport could be reclaimed. They were desperate.

The robots were programmed with the skill levels of the sporting heroes on whom they had been modelled. The message was clear: no more cheating.

The media weren't the only ones pleased by this solution. Sports veterans (and their dwindling bank accounts) were glad to